OM nucleic - nucleic search, using sw model

Copyright   GenCore version 5.1.6
            Copyright (c) 1993 - 2004 Compugen Ltd.

Run on: October 8, 2004, 04:42:22 ; Search time 8348 Seconds
            (without alignments)
            11183.624 Million cell updates/sec

Title: US-10-086-156-23
Perfect score: 2154
Sequence: 1 atgacgatggcggtttgcg........taataaaaaaaaaaaaaa 2154

Scoring table: IDENTITY_NUC
                Gapop 10.0 , Gapext 1.0

Searched: 3470272 seqs, 2167516995 residues

Total number of hits satisfying chosen parameters: 6940544

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
                 Maximum Match 100%
                 Listing first 45 summaries

Database :   GenEmbl:*
        1:   gb_ba:*
        2:   gb_htg:*
        3:   gb_in:*
        4:   gb_om:*
        5:   gb_ov:*
        6:   gb_pat:*
        7:   gb_ph:*
        8:   gb_pl:*
        9:   gb_pr:*
       10:   gb_ro:*
       11:   gb_sts:*
       12:   gb_sy:*
       13:   gb_un:*
       14:   gb_vi:*
       15:   gb_ba:*
       16:   em_fun:*
       17:   em_hum:*
       18:   em_in:*
       19:   em_mu:*
       20:   em_om:*
       21:   em_or:*
       22:   em_ov:*
       23:   em_pat:*
       24:   em_ph:*
       25:   em_pl:*
       26:   em_ro:*
       27:   em_sts:*
       28:   em_un:*
       29:   em_vi:*
       30:   em_htg_hum:*
       31:   em_htg_inv:*
       32:   em_htg_other:*
       33:   em_htg_mus:*
       34:   em_htg_pln:*
       35:   em_htg_rod:*
       36:   em_htg_mam:*
       37:   em_htg_vrt:*
       38:   em_sy:*
       39:   em_htgo_hum:*
       40:   em_htgo_mus:*
       41:   em_htgo_other:*

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
c 1	1200.4	55.7	135044	9	AC006001	AC006001 Homo sapi
2	1160.6	53.9	194464	2	AC146119	AC146119 Pan trogl
3	1031.4	47.9	135044	9	AC006001	AC006001 Homo sapi
4	588.6	27.3	4807	6	BD183414	BD183414 Novel gen
5	587	27.3	1211	9	BC042482	BC042482 Homo sapi
6	587	27.3	2576	6	AX714361	AX714361 Sequence
7	587	27.3	2576	6	AK056631	AK056631 Homo sapi
8	565.2	26.2	1124	6	BD275557	BD275557 MOLECULES
9	433.6	20.1	3864	6	AB056802	AB056802 Macaca fa
10	387.8	18.0	680	6	AX575790	AX575790 Sequence
c 11	374	17.4	374	11	G51111	G51111 SHGC-80071
12	315.2	14.6	363	10	AX575786	AX575786 Sequence
c 13	293.6	13.6	188791	10	AC117185	AC117185 Mus muscu
c 14	293.6	13.6	257003	10	AC122830	AC122830 Mus muscu
15	268.8	12.5	1103	10	BC051544	BC051544 Mus muscu
16	257.4	11.9	4143	9	AK127790	AK127790 Homo sapi
17	217.6	10.1	667	11	BV029947	BV029947 S212P6093
18	214.8	10.0	246	6	AX887431	AX887431 Sequence
19	214.8	10.0	246	6	BD027041	BD027041 Sequence
c 20	177.4	8.2	194464	2	AC146119	AC146119 Pan trogl
c 21	172.4	8.0	321	6	AX575788	AX575788 Sequence
c 22	167.2	7.8	203685	10	AC122339	AC122339 Mus muscu
c 23	159.2	7.4	213729	2	AC116246	AC116246 Rattus no
24	159.2	7.4	216180	2	AC119707	AC119707 Rattus no
25	121.2	5.6	178538	2	AC134778	AC134778 Homo sapi
26	121.2	5.6	203192	9	AC103810	AC103810 Homo sapi
27	118.4	5.5	79776	2	AC139431	AC139431 Homo sapi
28	117.8	5.5	82659	2	AC135725	AC135725 Homo sapi
29	117.8	5.5	204940	2	AC069007	AC069007 Homo sapi
c 30	117.8	5.5	207611	9	AC091132	AC091132 Homo sapi
31	116.4	5.4	18534	9	HSSERCA1	Y15724 Homo sapien
c 32	116.4	5.4	111370	9	AC067815	AC067815 Homo sapi
c 33	116.4	5.4	134362	9	AC084853	AC084853 Homo sapi
34	116.4	5.4	158414	2	AC005940	AC005940 Homo sapi
c 35	116.2	5.4	179564	9	AC139677	AC139677 Homo sapi
36	116.2	5.4	198470	9	AC046170	AC046170 Homo sapi
37	116.2	5.4	200420	9	AC138645	AC138645 Homo sapi
38	115.8	5.4	73947	9	AC068152	AC068152 Homo sapi
39	115.8	5.4	88004	9	AC002316	AC002316 Homo sapi
c 40	115.8	5.4	140386	9	AC015921	AC015921 Homo sapi
41	115.8	5.4	164874	2	AC138223	AC138223 Homo sapi
42	115.2	5.3	183093	9	AC138688	AC138688 Homo sapi
c 43	115	5.3	28746	2	AC008647	AC008647 Homo sapi
44	115	5.3	101744	9	AL512783	AL512783 Human DNA
45	114.6	5.3	122168	9	AC127383	AC127383 Homo sapi

ALIGNMENTS

RESULT 1
AC006001/c
LOCUS   AC006001   135044 bp   DNA   linear   PRI 02-OCT-2003
DEFINITION Homo sapiens PAC clone RP4-756H11 from 7, complete sequence.
ACCESSION   AC006001
VERSION   AC006001.2 GI:5708496
KEYWORDS   HTG.
SOURCE   Homo sapiens (human)
ORGANISM   Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1 (bases 1 to 135044)
AUTHORS   Sulston,J.E. and Wilson,R.
TITLE   Toward a complete human genome sequence
JOURNAL   Genome Res. 8 (11), 1097-1108 (1998)

Pred. No. is the number of results predicted by chance to have a

```
MEDLINE     99063792
PUBMED      9847074
REFERENCE   2  (bases 1 to 135044)
AUTHORS     Lamar,B., Le.T. and Wohldmann,P.
TITLE       The sequence of Homo sapiens PAC clone RP4-756H11
JOURNAL     Unpublished (2001)
REFERENCE   3  (bases 1 to 135044)
AUTHORS     Waterston,R.H.
TITLE       Direct Submission
JOURNAL     Submitted (22-NOV-1998) Genome Sequencing Center, Washington
            University School of Medicine, 4444 Forest Park Parkway, St. Louis,
            MO 63108, USA
REFERENCE   4  (bases 1 to 135044)
AUTHORS     Waterston,R.H.
TITLE       Direct Submission
JOURNAL     Submitted (07-AUG-1999) Genome Sequencing Center, Washington
            University School of Medicine, 4444 Forest Park Parkway, St. Louis,
            MO 63108, USA
REFERENCE   5  (bases 1 to 135044)
AUTHORS     Waterston,R.
TITLE       Direct Submission
JOURNAL     Submitted (30-SEP-2000) Department of Genetics, Washington
            University, 4444 Forest Park Avenue, St. Louis, Missouri 63108, USA
REFERENCE   6  (bases 1 to 135044)
AUTHORS     Waterston,R.
TITLE       Direct Submission
JOURNAL     Submitted (26-APR-2003) Department of Genetics, Washington
            University, 4444 Forest Park Avenue, St. Louis, Missouri 63108, USA
REFERENCE   7  (bases 1 to 135044)
AUTHORS     Wilson,R.
TITLE       Direct Submission
JOURNAL     Submitted (02-OCT-2003) Department of Genetics, Washington
            University, 4444 Forest Park Avenue, St. Louis, Missouri 63108, USA
COMMENT     On Aug 8, 1999 this sequence version replaced gi:3907522.
            -------------- Genome Center
            Center: Washington University Genome Sequencing Center
            Center code: WUGSC
            Web site: http://genome.wustl.edu
            Contact: sapiens@watson.wustl.edu
            -------------- Summary Statistics
            Center project name: H_DJ0756H11
            --------------

            NOTICE: This sequence may not represent the entire insert of this
            clone. It may be shorter because we only sequence overlapping
            clone sections once, or longer because we provide a small overlap
            between neighboring data submissions.

            This sequence was finished as follows unless otherwise noted:
            all regions were double stranded, sequenced with an alternate
            chemistry, or covered by high quality data (i.e., phred quality >=
            30); an attempt was made to resolve all sequencing problems, such
            as compressions and repeats; all regions were covered by sequence
            from more than one subclone; and the assembly was confirmed by
            restriction digest.

            MAPPING INFORMATION:
            The sequence of this clone was established as part of a mapping and
            sequencing collaboration between the NHGRI Chromosome 7 Mapping
            Project (Eric D. Green, Director), John D. McPherson in the
            Department of Genetics (Washington University), and the Washington
            University Genome Sequencing Center. For additional information
            about the map position of this sequence, see
            http://www.nhgri.nih.gov/DIR/GTB/CHR7, send
            mailto:egreen@nhgri.nih.gov, or see http://genome.wustl.edu.

            SOURCE INFORMATION:
            This clone was derived from human PAC library RPCI-4, prepared by
            Pieter de Jong and coworkers at http://www.chori.org using the
            method described by Ioannou et al., Nature Genetics 6:84-9 (1994).
            The library is from one male donor.
            The clone may be obtained either from Genome Systems, Inc.
            (http://www.genomesystems.com) or Research Genetics, Inc.
            (http://www.resgen.com); or from Pieter de Jong.
            VECTOR: pCYPAC2

            NEIGHBORING SEQUENCE INFORMATION:
            Actual start of this clone is at base position 1 of RP4-756H11.
            actual end is at base position 135044 of RP4-756H11.
FEATURES             Location/Qualifiers
     source          1..135044
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9606"
                     /chromosome="7"
                     /map="7"
                     /clone="RP4-756H11"
                     /clone_lib="RPCI-4"
     repeat_region   59..349
                     /rpt_family="Alu"
     repeat_region   369..488
                     /rpt_family="ERV1"
     repeat_region   494..800
                     /rpt_family="Alu"
     repeat_region   835..1030
                     /rpt_family="L2"
     repeat_region   1357..1436
                     /rpt_family="Alu"
     repeat_region   1523..1828
                     /rpt_family="Alu"
     repeat_region   2038..2314
                     /rpt_family="Alu"
     repeat_region   2328..2631
                     /rpt_family="Alu"
     repeat_region   2686..2781
                     /rpt_family="Alu"
     repeat_region   2785..2927
                     /rpt_family="L1"
     repeat_region   2937..2973
                     /rpt_family="L1"
     repeat_region   2974..3228
                     /rpt_family="L1"
     repeat_region   3229..3516
                     /rpt_family="Alu"
     repeat_region   3517..3546
                     /rpt_family="L1"
     repeat_region   3547..3793
                     /rpt_family="MER2_type"
     repeat_region   3794..3823
                     /rpt_family="L1"
     repeat_region   3824..4073
                     /rpt_family="Alu"
     repeat_region   4074..4144
                     /rpt_family="L1"
     repeat_region   4156..4643
                     /rpt_family="L1"
     repeat_region   4657..4887
                     /rpt_family="Alu"
     repeat_region   4895..5060
                     /rpt_family="MIR"
     repeat_region   5162..5624
                     /rpt_family="L1"
     repeat_region   5625..5919
                     /rpt_family="Alu"
     repeat_region   5920..6087
                     /rpt_family="L1"
     repeat_region   6214..6370
                     /rpt_family="Alu"
     repeat_region   6492..6700
                     /rpt_family="MIR"
     repeat_region   7521..7616
                     /rpt_family="MIR"
     repeat_region   7652..7971
                     /rpt_family="Alu"
     repeat_region   8017..8294
                     /rpt_family="Alu"
```

```
repeat_region   8299..8437
    /rpt_family="Alu"
repeat_region   8769..8930
    /rpt_family="Alu"
-               8972..9065
    /rpt_family="Alu"
repeat_region   9079..9126
    /rpt_family="(TA)n"
repeat_region   9187..9229
    /rpt_family="(TG)n"
repeat_region   9230..9250
    /rpt_family="AT_rich"
repeat_region   9251..9549
    /rpt_family="Alu"
repeat_region   9550..9572
    /rpt_family="AT_rich"
repeat_region   9648..9678
    /rpt_family="(A)n"
repeat_region   9750..9791
    /rpt_family="Alu"
repeat_region   9808..10423
    /rpt_family="L2"
repeat_region   11152..11226
    /rpt_family="MIR"
repeat_region   12015..12333
    /rpt_family="Alu"
repeat_region   11716..12011
    /rpt_family="Alu"
repeat_region   12465..12583
    /rpt_family="MIR"
repeat_region   13722..13853
    /rpt_family="MER1_type"
repeat_region   14062..14287
    /rpt_family="L2"
repeat_region   14461..14592
    /rpt_family="MIR"
repeat_region   14669..14965
    /rpt_family="Alu"
repeat_region   15509..15814
    /rpt_family="Alu"
repeat_region   15827..16007
```

```
Query Match      55.7%;   Score 1200.4;   DB 9;   Length 135044;
Best Local Similarity 99.8%;   Pred. No. 1.8e-272;
Matches 1212;   Conservative 0;   Mismatches 0;   Indels 1;   Gaps 1;

Qy    926  AAAGGCGGCTGTATTGGAAACTCCATTGGTCCAGAGAGACAGAACAATGAGAAGAAAT  985
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  32238  AAAGGCGGCTGTATTGGAAACTCCATTGGTCCAGAGAGACAGAACAATGAGAAGAAAT  32179

Qy    986  CCCCTGTCCAGTTGCCTGCGAGGAGTATTCCAACACTTCATGGGCTAGAGGATTCCATTGA  1045
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  32178  CCCCTGTCCAGTTGCCTGCGAGGAGTATTCCAACACTTCATGGGCTAGAGGATTCCATTGA  32119

Qy   1046  GATGGGGTTTACGTCTTGATTTTGAAACACCTGTCAGCACTGTTCTCTGTTTGCATGGCAA  1105
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  32118  GATGGGGTTTACGTCTTGATTTTGAAACACCTGTCAGCACTGTTCTCTGTTTGCATGGCAA  32059

Qy   1106  TTCTGACCCTTTATGGCAACAACACACCCCTGGGACAACCCAGATTTGTAGATTGAGATCC  1165
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  32058  TTCTGACCCTTTATGGCAACAACACACCCCTGGGACAACCCAGATTTGTAGATTGAGATCC  31999

Qy   1166  AAAGGTAGAATTTCCAGACAGCAGTCCAACCAAGGTATCAAGTGATGTTCCAGAGTGGAAGG  1225
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  31998  AAAGGTAGAATTTCCAGACAGCAGTCCAACCAAGGTATCAAGTGATGTTCCAGAGTGGAAGG  31939

Qy   1226  CTCTCACCGTGTCCCAGGATTTCTGGGGTTTGTAAGCAGTACTGGCCATTTGTTGACCCTG  1285
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  31938  CTCTCACCGTGTCCCAGGATTTCTGGGGTTTGTAAGCAGTACTGGCCATTTGTTGACCCTG  31879

Qy   1286  TTTTTTTACCTAATCATTCGTGTCTTTTTAGGACATGGTTTTACCCGATCCCTGCCAAGGA  1345
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  31878  TTTTTTTACCTAATCATTCGTGTCTTTTTAGGACATGGTTTTACCCGATCCCTGCCAAGGA  31819
```

```
Qy   1346  TCCAGAATTCCAATAGCTGAAAACCCTGTATAGCTTTTCTCCTATTCTGCCTTACCCAA  1405
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  31818  TCCAGAATTCCAATAGCTGAAAACCCTGTATAGCTTTTCTCCTATTCTGCCTTACCCAA  31759

Qy   1406  GACACACTTGAACCCTCAGTAAGGCTATAGAGAGGGCCATGAGCAGGCAGCCTCTCC  1465
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  31758  GACACACTTGAACCCTCAGTAAGGCTATAGAGAGGGCCATGAGCAGGCAGCCTCTCC  31699

Qy   1466  CTTGTTCTCACAGCTCAGATGAGGGGTTGACTGAGGCCAGCAATCCTTGTAGGTGTGA  1525
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  31698  CTTGTTCTCACAGCTCAGATGAGGGGTTGACTGAGGCCAGCAATCCTTGTAGGTGTGA  31639

Qy   1526  CAGTTGCAATATAATTAACAGTTCAAGATCTAGAAGTACCTTTTGAAAGAACCCCTTCA  1585
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  31638  CAGTTGCAATATAATTAACAGTTCAAGATCTAGAAGTACCTTTTGAAAGAACCCCTTCA  31579

Qy   1586  GGGATATCTATCCACAGTAGCCTGGACGCAGCCAAGGTGAACCTGAGACATTTGACCCACAC  1645
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  31578  GGGATATCTATCCACAGTAGCCTGGACGCAGCCAAGGTGAACCTGAGACATTTGACCCACAC  31519

Qy   1646  AATAAGGAGGGGGCCATTCTTCTTTTCAAATATTTGGCTTCAGAATACACTTCATTACACAT  1705
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  31578  AATAAGGAGGGGGCCATTCTTCTTTTCAAATATTTGGCTTCAGAATACACTTCATTACACAT  31518

Qy   1706  GCAAATATTGAGAGATTAACAGAAATTCCAGCTCTTATGCCTAACTGAGAAGAGCCACTG  1765
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  31458  GCAAATATTGAGAGATTAACAGAAATTCCAGCTCTTATGCCTAACTGAGAAGAGCCACTG  31399

Qy   1766  CAAGTTGCAGTTAGGTACCCATGTGCAGCAGAGGCCAGCTGAATCCCAGAGCTTCCCAAA  1825
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  31458  CAAGTTGCAGTTAGGTACCCATGTGCAGCAGAGGCCAGCTGAATCCCAGAGCTTCCCAAA  31398

Qy   1826  GTGGACACCAGCGGGGACTATTCCTGATGTCCCACCAGAGAGGAAGATGAGCTGAGGC  1885
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  31338  GTGGACACCAGCGGGGACTATTCCTGATGTCCCACCAGAGAGGAAGATGAGCTGAGGC  31279

Qy   1886  GCTCTTGCTTCTGCCCAAATGCATCCCATGTGCATTCACGTGTCACCCATTCAAAATAACA  1945
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  31278  GCTGTTGCTTCTGCCCAAATGCATCCCATGTGCATTCACGTGTCACCCATTCAAAATAACA  31219

Qy   1946  TGGGCATTCTTGGAACCTTGTGATCTGACATGTGAAGACCAGCCTACACCATTGGGGTGC  2005
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  31218  TGGGCATTCTTGGAACCTTGTGATCTGACATGTGAAGACCAGCCTACACCATTGGGGTGC  31159

Qy   2006  AGGGGCTCACACTTGTAATCCTAGCACTTTGGAAGGCTGGAGGTGGGCAGATTGCTTGAGC  2065
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  31158  AGGGGCTCACACTTGTAATCCTAGCACTTTGGAAGGCTGGAGGTGGGCAGATTGCTTGAGC  31099

Qy   2066  ACAGGAGTTCCAGACCAGCCTGAGCAACATGGCGAAATCCTGTCTTCTTCAAGAAAATAA-A  2124
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  31098  ACAGGAGTTCCAGACCAGCCTGAGCAACATGGCGAAATCCTGTCTTCTTCAAGAAAATAATA  31039

Qy   2125  ATAATAATAATAAT  2138
           ||||||||||||||
Db  31038  ATAATAATAATAAT  31025
```

```
RESULT 2
AC146119
LOCUS       AC146119   194464 bp   DNA   linear   HTG   01-AUG-2003
DEFINITION  Pan troglodytes chromosome UNK clone RP43-38J3, *** SEQUENCING IN
            PROGRESS ***, 31 unordered pieces.
ACCESSION   AC146119
VERSION     AC146119.1   GI:33387073
KEYWORDS    HTG; HTGS_PHASE1.
SOURCE      Pan troglodytes (chimpanzee)
  ORGANISM  Pan troglodytes
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Pan.
REFERENCE   1  (bases 1 to 194464)
  AUTHORS   Wilson,R.K.
  TITLE     The sequence of Pan troglodytes clone
```

```
JOURNAL     Unpublished
REFERENCE   2 (bases 1 to 194464)
AUTHORS     Wilson.R.K.
TITLE       Direct Submission
JOURNAL     Submitted (01-AUG-2003) Genetics, Genome Sequencing Center, 4444
            Forest Park Parkway, St. Louis, MO 63108, USA
COMMENT     -------- Genome Center --------
            Center: Washington University Genome Sequencing Center
            Center code: WUGSC
            Web site:http://genome.wustl.edu/gsc/index.shtml
            -------- Project Information --------
            Center project name: C_PT038J03
            -------- Summary Statistics --------
            Sequencing vector: M13; 0%
            Sequencing vector: plasmid; 100%
            Chemistry: Dye-primer ET; 0% of reads
            Chemistry: Dye-terminator Big Dye; 100% of reads
            Assembly program: Phrap; version 0.990319
            Consensus quality: 177293 bases at least Q40
            Consensus quality: 179213 bases at least Q30
            Consensus quality: 180602 bases at least Q20
          * NOTE: This is a 'working draft' sequence. It currently
          * consists of 31 contigs. The true order of the pieces
          * is not known and their order in this sequence record is
          * arbitrary. Gaps between the contigs are represented as
          * runs of N, but the exact sizes of the gaps are unknown.
          * This record will be updated with the finished sequence
          * as soon as it is available and the accession number will
          * be preserved.
          *      1    1329: contig of 1329 bp in length
          *   1330    1429: gap of unknown length
          *   1430    3053: contig of 1624 bp in length
          *   3054    3153: gap of unknown length
          *   3154    4975: contig of 1822 bp in length
          *   4976    5075: gap of unknown length
          *   5076    6209: contig of 1134 bp in length
          *   6210    6309: gap of unknown length
          *   6310    7326: contig of 1017 bp in length
          *   7327    7426: gap of unknown length
          *   7427   10326: contig of 2900 bp in length
          *  10327   10426: gap of unknown length
          *  10427   12412: contig of 1986 bp in length
          *  12413   12512: gap of unknown length
          *  12513   14266: contig of 1754 bp in length
          *  14267   14366: gap of unknown length
          *  14367   16448: contig of 2082 bp in length
          *  16449   16548: gap of unknown length
          *  16549   18757: contig of 2209 bp in length
          *  18758   18857: gap of unknown length
          *  18858   22270: contig of 3413 bp in length
          *  22271   22370: gap of unknown length
          *  22371   25054: contig of 2684 bp in length
          *  25055   25154: gap of unknown length
          *  25155   28295: contig of 3141 bp in length
          *  28296   28395: gap of unknown length
          *  28396   31482: contig of 3087 bp in length
          *  31483   31582: gap of unknown length
          *  31583   34542: contig of 2960 bp in length
          *  34543   34642: gap of unknown length
          *  34643   37879: contig of 3237 bp in length
          *  37880   37979: gap of unknown length
          *  37980   43297: contig of 5318 bp in length
          *  43298   43397: gap of unknown length
          *  43398   48231: contig of 4834 bp in length
          *  48232   48331: gap of unknown length
          *  48332   55882: contig of 7551 bp in length
          *  55883   55982: gap of unknown length
          *  55983   64790: contig of 8808 bp in length
          *  64791   64890: gap of unknown length
          *  64891   73831: contig of 8941 bp in length
          *  73832   73931: gap of unknown length
          *  73932   82214: contig of 8383 bp in length
          *  82215   82314: gap of unknown length
          *  82315   90706: contig of 8292 bp in length
          *  90707   90806: gap of unknown length
          *  90807   98797: contig of 7991 bp in length
          *  98798   98897: gap of unknown length
          *  98898  106699: contig of 7802 bp in length
          * 106700  106799: gap of unknown length
          * 106800  117151: contig of 10352 bp in length
          * 117152  117251: gap of unknown length
          * 117252  129326: contig of 12075 bp in length
          * 129327  129426: gap of unknown length
          * 129427  145201: contig of 15775 bp in length
          * 145202  145301: gap of unknown length
          * 145302  161680: contig of 16379 bp in length
          * 161681  161780: gap of unknown length
          * 161781  193229: contig of 31449 bp in length
          * 193230  193329: gap of unknown length
          * 193330  194464: contig of 1135 bp in length.
FEATURES             Location/Qualifiers
     source          1..194464
                     /organism="Pan troglodytes"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9958"
                     /chromosome="UNK"
                     /clone="RP43-38J3"
     misc_feature    1..1329
                     /note="assembly_name:Contig10"
     misc_feature    1430..3053
                     /note="assembly_name:Contig11"
     misc_feature    3154..4975
                     /note="assembly_name:Contig17"
     misc_feature    5076..6209
                     /note="assembly_name:Contig18"
     misc_feature    6310..7326
                     /note="assembly_name:Contig19"
     misc_feature    7427..10326
                     /note="assembly_name:Contig20"
     misc_feature    10427..12412
                     /note="assembly_name:Contig21"
     misc_feature    12513..14266
                     /note="assembly_name:Contig22"
     misc_feature    14367..16448
                     /note="assembly_name:Contig23"
     misc_feature    16549..18757
                     /note="assembly_name:Contig24"
     misc_feature    18858..22270
                     /note="assembly_name:Contig25"
     misc_feature    22371..25054
                     /note="assembly_name:Contig26"
     misc_feature    25155..28295
                     /note="assembly_name:Contig27"
     misc_feature    28396..31482
                     /note="assembly_name:Contig28"
     misc_feature    31583..34542
                     /note="assembly_name:Contig29"
     misc_feature    34643..37879
                     /note="assembly_name:Contig30"
     misc_feature    37980..43297
                     /note="assembly_name:Contig31"
     misc_feature    43398..48231
                     /note="assembly_name:Contig32"
     misc_feature    48332..55882
                     /note="assembly_name:Contig33"
     misc_feature    55983..64790
                     /note="assembly_name:Contig34"
     misc_feature    64891..73831
                     /note="assembly_name:Contig35"
     misc_feature    73932..82314
                     /note="assembly_name:Contig36"
     misc_feature    82415..90706
                     /note="assembly_name:Contig37"
```

```
misc_feature    90807..98797
                /note="assembly_name:Contig38"
misc_feature    98898..106699
                /note="assembly_name:Contig39"
misc_feature    106800..117151
                /note="assembly_name:Contig40"
misc_feature    117252..129326
                /note="assembly_name:Contig41"
misc_feature    129427..145201
                /note="assembly_name:Contig42"
misc_feature    145302..161680
                /note="assembly_name:Contig43"
misc_feature    161781..193229
                /note="assembly_name:Contig44"
misc_feature    193330..194464
                /note="assembly_name:Contig6"
ORIGIN

Query Match     53.9%;  Score 1160.6;  DB 2;  Length 194464;
Best Local Similarity   98.7%;  Pred. No. 4.8e-263;
Matches 1203;  Conservative   0;  Mismatches   9;  Indels   7;  Gaps   3;

Qy    926 AAAGGCGGCTGTATTTGGAAACTCCCATTGGTCCAGAGAGACAGAACAATGAGAAGAAAT 985
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 142260 AAAGGCGGCTGTATTTGGAAACTCCCATTGGTCCAGAGAGACAGAACAATGAGAAGAAAT 142319

Qy    986 CCCCTGTCCAGTTGCCTGCGAGGAGTATTCCAACACTTCATGGGCTAGAGGATTCCATTGA 1045
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 142320 CCCCTGTCCAGTTGCCTGCGAGGAGTATTCCAGCACTTCATGGGCTAGAGGATTCCATTGA 142379

Qy   1046 GATGGGGTTTACGTCTTGATTTGAACACTGTTCTCGTTTGTTTGCATGGCAA 1105
          |||||||||||||||||||||||||||||||||||||||| ||||||||
Db 142380 GATGGGGTTTACGTCTTGATTTGAACACTGTTCTC----TGCATGGCAA 142435

Qy   1106 TTCTGACCCTTTATGGCAACAACACCCCTGGGACAACAACCCAGATTTGTAGATTGAGATCC 1165
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 142436 TTCTGACCCTTTATGGCAACAACACCCCTGGGACAACAACCCAGATTTGTAGATTGAGATCC 142495

Qy   1166 AAAGGTAGAATTCCAGACAGTCCAACCAAGGTATCAAGTGATGTTCCAGAGTGGAAGG 1225
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 142496 AAAGGTAGAATTCCAGACAGTCCAACCAAGGTATCAAGTGATGTTCCAGAGTGGAAGG 142555

Qy   1226 CTCTCACCGTGTCCCAGGATTTCTGGGGTTTGTAAGCAGTACTGGCCATTTGTGACCCTG 1285
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 142556 CTCTCACCATGTCCCAGGATTTCTGGGGTTTGTAAGCAGTACTGGCCATTTGTGACCCTG 142615

Qy   1286 TTTTTTACCTAATCATTCTGTCTTTTTAGGACATGGTTTTACCCGATCCCTGGCAAAGGA 1345
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 142616 TTTTTTACCTAATCATTCTGTCTTTTTAGGACATGGTTTTACCCGATCCCTGGCAAAGGA 142675

Qy   1346 TCCAGAATTCCAATAGCTGAAAACCCTGTATAGCTTTTCTCCTATTCGCCTTACCCAA 1405
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 142676 TCCAGAATTCCAATAGCTGAAAACCCTGTATAGCTTTTCTCCTATTCGCCTTACCCAA 142735

Qy   1406 GACACACTTGAACCCCTCAGTAAGGCTATAGAGAGAGGGCCATGAGCAGGGGCAGCCTCTCC 1465
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 142736 GACACACTTGAACCCCTCAGTAAGGCTATAGAGAGAGGGCCATGAGCAGGGGCAGCCTCTCC 142795

Qy   1466 CTTGTTTCTACAGTCCATGATGAGGGGTTGACTGAGGCCAGCAATCCTTGTAGGTGTGA 1525
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 142796 CTTATTTCTACAGTCCATGATGAGGGGTTGACTGAGGCCAGCAATCCTTGTAGGTGTGA 142855

Qy   1526 CAGTTGCAATATAATTAACAGTTTCAAGATCTAGAGGTACCTTTGAAAGAACCCCTTCA 1585
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 142856 CAGTTGCAATATAATTAACAGTTTCAAGATCTAGAGGTACCTTTGAAAGAACCCCTTCA 142915

Qy   1586 GGGATATCTATCCACAGTAGCCTGGAGCAGCCAAGGTGAACCTGAGATTTTGACCCACAC 1645
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 142916 GGGATATCTATCCACAGTAGCCTGGAGCAGCCAAAGTGAACCTGAGATTTTGACCCACAC 142975

Qy   1646 AATAAGGGGGGCCATTCTTTTTCAAATATTTTGGCTTCAGAATACACTTCATTACACAT 1705
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 142976 AATAAGGGGGGCCATTCTTTTTCAAATATTTTGGCTTCAGAATACACTTCATTACACAT 143035
```

```
Qy   1706 GCAAATATTGAGAGATTAACAGAAATTCCAGCTCTTATGCTAACTGAGAAGAGAGCCACTG 1765
          |||||||||||||||||||||||||||||| |||||||||||||||||||||||||||
Db 143036 GCAAATATTGAGAGATTAAC--AAATTCCAGCTCTTATGCTAACTGAGAAGAGAGCCACTG 143093

Qy   1766 CAAGTTGCAGTTAGGTACCCATGTGCAGAGGCCAGCGTGAATCCCAGAGCTTCCCAAA 1825
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 143094 GAAGTTGCAGTTAGGTACCCATGTGCAGAGGCCAGCGTGAATCCCAGAGCTTCCCAAA 143153

Qy   1826 GTGGACACCAGCGGGGACTATTCCTGATGTCCACCCAGAGAGAGAAGATGAGCTGAGGC 1885
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 143154 GTGGACACCAGCGGGGACTATTCCTGATGTCCACCCAGAGAGAGAAGATGAGCTGAGGC 143213

Qy   1886 GCTCTTGCTCTGCCCAAATGCATCCCATGTGCATTCACGTGTCACCCATTCAAAATAACA 1945
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 143214 GCTGTTGCTCTGCCCAAATGCATCCCATGTGCATTCACGTGTCACCCATTCAAAATAACA 143273

Qy   1946 TGGCATTCTTGGAACCTTGTATCTGACAACGTGAGCCTACACATTGGGTGGGTGC 2005
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 143274 TGGCATTCTTGGAACCTTGTATCTGACAACGTGAGCCTACACATTGGGTGGGTGC 143333

Qy   2006 AGGGGCTCACACTTGTAATCCTAGCACTTTGGAAGGCTGAGGTGGGCAGATTGCTTGAGC 2065
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 143334 AGTGGCTCACACTTGTAATCCTAGCACTTTGGAAGGCTGAGGTGGGCAGATTGCTTGAGC 143393

Qy   2066 ACAGGAGTTCCAGACCAGCCTGAGCAACATGGGCGAAATCCTGTCTTCTTCAAGAAATA-AA 2124
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 143394 ACAGGAGTTCCAGACCAGCCTGAGCAACATGGGCGAAATCCTGTCTTCTTCAAGAAATACAA 143453

Qy   2125 ATAATAATAATAATAAAA 2143
          |||||||||||||||||
Db 143454 ATAATAATAATAATAATAA 143472
```

```
RESULT 3
AC006001
LOCUS       AC006001              135044 bp    DNA     linear   PRI 02-OCT-2003
DEFINITION  Homo sapiens PAC clone RP4-756H11 from 7, complete sequence.
ACCESSION   AC006001
VERSION     AC006001.2  GI:5708496
KEYWORDS    HTG.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 135044)
  AUTHORS   Sulston,J.E. and Wilson,R.
  TITLE     Toward a complete human genome sequence
  JOURNAL   Genome Res. 8 (11), 1097-1108 (1998)
  MEDLINE   99063792
   PUBMED   9847074
REFERENCE   2  (bases 1 to 135044)
  AUTHORS   Lamar,B., Le,T. and Wohldmann,P.
  TITLE     The sequence of Homo sapiens PAC clone RP4-756H11
  JOURNAL   Unpublished (2001)
REFERENCE   3  (bases 1 to 135044)
  AUTHORS   Waterston,R.H.
  TITLE     Direct Submission
  JOURNAL   Submitted (22-NOV-1998) Genome Sequencing Center, Washington
            University School of Medicine, 4444 Forest Park Parkway, St. Louis,
            MO 63108, USA
REFERENCE   4  (bases 1 to 135044)
  AUTHORS   Waterston,R.H.
  TITLE     Direct Submission
  JOURNAL   Submitted (07-AUG-1999) Genome Sequencing Center, Washington
            University School of Medicine, 4444 Forest Park Parkway, St. Louis,
            MO 63108, USA
REFERENCE   5  (bases 1 to 135044)
  AUTHORS   Waterston,R.
  TITLE     Direct Submission
  JOURNAL   Submitted (30-SEP-2000) Department of Genetics, Washington
            University, 4444 Forest Park Avenue, St. Louis, Missouri 63108, USA
REFERENCE   6  (bases 1 to 135044)
```

```
AUTHORS   Waterston,R.
TITLE     Direct Submission
JOURNAL   Submitted (26-APR-2003) Department of Genetics, Washington
          University, 4444 Forest Park Avenue, St. Louis, Missouri 63108, USA
REFERENCE 7  (bases 1 to 135044)
AUTHORS   Wilson,R.
TITLE     Direct Submission
JOURNAL   Submitted (02-OCT-2003) Department of Genetics, Washington
          University, 4444 Forest Park Avenue, St. Louis, Missouri 63108, USA
COMMENT   On Aug 8, 1999 this sequence version replaced gi:3907522.
          ------------ Genome Center
             Center: Washington University Genome Sequencing Center
             Center code: WUGSC
             Web site: http://genome.wustl.edu
             Contact: sapiens@watson.wustl.edu
          ------------ Summary Statistics
             Center project name: H_DJ0756H11
          ------------

          NOTICE:  This sequence may not represent the entire insert of this
          clone.  It may be shorter because we only sequence overlapping
          clone sections once, or longer because we provide a small overlap
          between neighboring data submissions.

          This sequence was finished as follows unless otherwise noted:
          all regions were double stranded, sequenced with an alternate
          chemistry, or covered by high quality data (i.e., phred quality >=
          30); an attempt was made to resolve all sequencing problems, such
          as compressions and repeats; all regions were covered by sequence
          from more than one subclone; and the assembly was confirmed by
          restriction digest.

          MAPPING INFORMATION:
          The sequence of this clone was established as part of a mapping and
          sequencing collaboration between the NHGRI Chromosome 7 Mapping
          Project (Eric D. Green, Director), John D. McPherson in the
          Department of Genetics (Washington University), and the Washington
          University Genome Sequencing Center. For additional information
          about the map position of this sequence, see
          http://www.nhgri.nih.gov/DIR/GTB/CHR7, send
          mailto:egreen@nhgri.nih.gov, or see http://genome.wustl.edu

          SOURCE INFORMATION:
          This clone was derived from human PAC library RPCI-4, prepared by
          Pieter de Jong and coworkers at http://www.chori.org using the
          method described by Ioannou et al., Nature Genetics 6:84-9 (1994).
          The library is from one male donor.
          The clone may be obtained either from Genome Systems, Inc.
          (http://www.genomesystems.com) or Research Genetics, Inc.
          (http://www.reggen.com); or from Pieter de Jong.
          VECTOR: pCYPAC2

          NEIGHBORING SEQUENCE INFORMATION:
          Actual start of this clone is at base position 1 of RP4-756H11.
          actual end is at base position 135044 of RP4-756H11.

FEATURES          Location/Qualifiers
     source        1..135044
                   /organism="Homo sapiens"
                   /mol_type="genomic DNA"
                   /db_xref="taxon:9606"
                   /chromosome="7"
                   /map="7"
                   /clone="RP4-756H11"
                   /clone_lib="RPCI-4"
     repeat_region 59..349
                   /rpt_family="Alu"
     repeat_region 369..488
                   /rpt_family="ERV1"
     repeat_region 494..800
                   /rpt_family="Alu"
     repeat_region 835..1030
                   /rpt_family="L2"
     repeat_region 1357..1436
```

```
     repeat_region 1523..1828
                   /rpt_family="Alu"
     repeat_region 2038..2314
                   /rpt_family="Alu"
     repeat_region 2328..2631
                   /rpt_family="Alu"
     repeat_region 2686..2781
                   /rpt_family="Alu"
     repeat_region 2785..2927
                   /rpt_family="Alu"
     repeat_region 2937..2973
                   /rpt_family="L1"
     repeat_region 2974..3228
                   /rpt_family="L1"
     repeat_region 3229..3516
                   /rpt_family="L1"
     repeat_region 3517..3546
                   /rpt_family="Alu"
     repeat_region 3547..3793
                   /rpt_family="L1"
     repeat_region 3794..3823
                   /rpt_family="MER2_type"
     repeat_region 3824..4073
                   /rpt_family="L1"
     repeat_region 4074..4144
                   /rpt_family="Alu"
     repeat_region 4156..4643
                   /rpt_family="L1"
     repeat_region 4657..4987
                   /rpt_family="L1"
     repeat_region 4895..5060
                   /rpt_family="Alu"
     repeat_region 5162..5524
                   /rpt_family="MIR"
     repeat_region 5625..5919
                   /rpt_family="L1"
     repeat_region 5920..6087
                   /rpt_family="Alu"
     repeat_region 6214..6370
                   /rpt_family="L1"
     repeat_region 6492..6700
                   /rpt_family="Alu"
     repeat_region 7521..7616
                   /rpt_family="MIR"
     repeat_region 7652..7971
                   /rpt_family="MIR"
     repeat_region 8017..8294
                   /rpt_family="Alu"
     repeat_region 8299..8437
                   /rpt_family="Alu"
     repeat_region 8769..8930
                   /rpt_family="Alu"
     repeat_region 8972..9065
                   /rpt_family="Alu"
     repeat_region 9079..9126
                   /rpt_family="Alu"
     repeat_region 9187..9229
                   /rpt_family="(TA)n"
     repeat_region 9230..9250
                   /rpt_family="(TG)n"
     repeat_region 9251..9549
                   /rpt_family="AT_rich"
     repeat_region 9550..9572
                   /rpt_family="Alu"
     repeat_region 9648..9678
                   /rpt_family="AT_rich"
     repeat_region 9750..9791
                   /rpt_family="(A)n"
     repeat_region 9808..10423
                   /rpt_family="Alu"
     repeat_region 11152..11226
                   /rpt_family="L2"
                   /rpt_family="MIR"
```

```
repeat_region    11716..12011
                 /rpt_family="Alu"
repeat_region    12015..12333
                 /rpt_family="Alu"
repeat_region    12465..12583
                 /rpt_family="MIR"
repeat_region    13722..13853
                 /rpt_family="MER1_type"
repeat_region    14062..14287
                 /rpt_family="L2"
repeat_region    14461..14592
                 /rpt_family="MIR"
repeat_region    14669..14965
                 /rpt_family="Alu"
repeat_region    15509..15814
                 /rpt_family="Alu"
repeat_region    15827..16007

Query Match            47.9%;  Score 1031.4;  DB 9;  Length 135044;
Best Local Similarity   93.6%;  Pred. No. 1.5e-232;
Matches 1143;  Conservative   0;  Mismatches  66;  Indels  12;  Gaps   6;

Qy    926 AAAGGCGGCGTATTGGAAAACTCCATTGGTCCAGAGAGACAGAACAATGAGAGAAAT 985
          ||||| || ||||||||||||||||||||||||||||| ||||||||||||||||||
Db  82276 AAAGGCAGCTATATTTGGAAACTCCCATTGGTCCAGAGAGGCAGAATAATGAGAAGAAAT 82335

Qy    986 CCCCTGTCCAGTTGCCTGCGAGGAGTATTCCAACACTT-CATGGGCTAGAGGATTCCATTG 1044
          |||||||||||||||||||| |||||||||||||||| |||||||||||||||||||||
Db  82336 CCCCTGTCCAGTTGCCTGCGAGGAATATTCCACCACTTCCATGGGCTAGAGGATTCCACTG 82395

Qy   1045 AGATGGGGTTACGTCTTGATTTTGAACACCTGTCAGCACTGTTCTCGTTTGCATGGCA 1104
          ||||||||||||||||| ||||||||||||||||||||||||||||||||||||||||
Db  82396 AGATAGGGGTTACGACTTGATTTTGAACACCTGTCAACA-TATTCTCTGTTTGCATGGCA 82454

Qy   1105 ATTCTGACCCTTTATGGCAACAACACACCCCTGGGACAACCCAGATTTGTAGATTGAGATC 1164
          |||||||||||||||||||||||||||||||||||||||||||||| ||||||||||||||
Db  82455 ATTCTGACCGTTTTATGCAACAATATACCCCTGGGACAACCTGGATTTGTGGACTGAGATC 82514

Qy   1165 CAAAGGTAGAAATTTCCAGACAGTCCAACCAAGGTATCAAGTGATGTTTCCAGAGTGGAAG 1224
          |||||||||||| |||||||||||||||||||||| |||||||||||||||||||||||||
Db  82515 CAAAGGCAGCAGTAGAATTTCCAGACAGTCCAACCAA--TATCAAGTGATGTTTCCAGAGTGGAAG 82572

Qy   1225 GCTCTCACCGTGTCCCAGGATTTCTGGGGGTTGTAAGCAGTACTGGCCATTTGTGACCCT 1284
          |||||||||||||||||||||||||| |||||||||||||||||||||||||||||||||
Db  82573 GCTCTCACAGTGTCCCAGGATTTCTGGGGTTTATAAGCAGTACTGGCCATTTGTGACCCT 82632

Qy   1285 GTTTTTACCTAATCATTCTGTCTTTTAGGACATGGTTTACCGATCCCTGGCAAAGG 1344
          ||||||||||||||||||||||||||||||||||| |||||||||||||||||||||
Db  82633 GTTTTTACCTAATCATTCTGTCTTTTTAGGATATGGGTTTTACCAGATCCCTGCCAAGG 82692

Qy   1345 ATCCAGAATTCCAATAGCTGAAAACCCTGTAATAGCTTTTCTCCTATTCTGCCTACCCA 1404
          |||||||||||||||||||||||||||||||| |||||||||||||||||||||||||||
Db  82693 ATCCAGCATTCCAATAGCTGAAAACCCCTG-TATATCTTTTCTCCTATCCTGCCTTACCCA 82751

Qy   1405 AGACACACTTGAACCCCCAGTAAGGCTATAGAGAGGGCCATGAGCAGGGGCAGCGCTCTC 1464
          |||||||||||||||||||||||||| ||||||||||||||||||||||||||||||||||
Db  82752 AGACACACTTGAACCCCTCAGTAAGGCTATACAGAGGGGCCATAGCAGGGGCAGCGCCTCTC 82811

Qy   1465 CCTTGTTTCTACAGCTCCACAGTAGCCTGAGCGGGGTTGACTGAGGCCAGCAATCCTTGTAGGTGTG 1524
          |||||||||||||||||||||||||| |||||||||| ||||||||||||||||||||||
Db  82812 CCTTGTTTCTGCAGCTCCACAGTAGCGGGGTTGGCTGAGGCCAGCAATCCTTGTAGGTGTG 82871

Qy   1525 ACAGTTGCAATATATAATTAACAGTTTCAAGATCCATCAGAGGTACCCTTTGAAAGAACCCCCTTC 1584
          |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  82872 ACAGTTTGCAATATATAATCAACAGTTTCAAGGTTTCAAGGTCTTACAGGCACCCGCTTTT 82931

Qy   1585 AGGGATAATCTATCCACAGTAGCCTGAGCAGCCAAGGCCAAGGCCAGCCACCA 1644
          |||||||||| ||||||||||| |||||||||||||||||||||||||||||||
Db  82932 AGGGATGCCTATCTACAGCCTGCTGGAGCAGCCAAGGTGAACCTGGAGATTTTGACCCACA 82991

Qy   1645 CAATAAGG--GGGGGCCATTCTTTTCAAATATTTGGCTTCAGAATACACTTCATTAC 1701
          |||||||   ||||||||||||| |||||||||||||||||||||||||||||||||
Db  82992 CAATAAGGGGTGGGGGCATTTTTTCAAATATTTCAAATATTTGGCTTCAGCAGACACTTCATTAC 83051
```

```
Qy   1702 ACATGCAAATATTGAGAGATTAACAGAAATTCCAGCTCTTTATGCCTAACTGAGAGAGCC 1761
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  83052 ATATGCAAATATTGAGAGATTAACAGAAATTCCAGCTCTTTATGCCTGACTCAGAAGAGCC 83111

Qy   1762 ACTGCAAGTTGCAGTTAGGTACCCATGTGCAGCAGAGGCCAGCCAGCTGAATCCCAGAGCTTCC 1821
          |||||||||||||||||||||||||||| |||||||||||||||||||||||||||||||||
Db  83112 ACTGCAAGTTGCAGTTAGGTACCCATGTGCAGCAGAGGCCAGCCAGCTGAATCCCAGAGCTTCC 83171

Qy   1822 CAAAGTGGACACCACGCGGGACTATTCCTGATGTCCACCCAAGAGAGGAAGATGAGCTG 1881
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||  ||
Db  83172 CAAAGTGGACACCACCAGCGGGACTATTCCTGATGTCCACCCAAGAGAGGAAGATGAACTC 83231

Qy   1882 AGGCGGCTCTTGCTCTGCCCAAATGCATCCCATGTGCATTCACGTGTCACCCATTCA---- 1937
          |||||||||||||||| ||||||||||||||||||||| ||||||||||||||| | ||
Db  83232 TGGTGCTGCTCTGCTCTGCCCAAATGCATCCCATGTGCATTCACGTGTCACCCATTCAAAAT 83291

Qy   1938 AAATAACATGGCATTCTTGGAACCTGTATCTGACATGTAAGCACCGCTACAACATTGGG 1997
          |||||||||||||||||||||||||||||||||||||||||||||||| ||||||||||
Db  83292 AAATAACATGGAATTCTTGGAACCTTGTATCTGACATGTAAGCACCAGCTTACACATTGGG 83351

Qy   1998 GTGGGTGCAGGGGCTCACACTTGTAATCCTAGCACTTTGGAAGGCTGAGGTGGGCAGATT 2057
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  83352 CTGGGTGCAGGGGCTCACACCTGTAATCCTAGCACACTTTGGAAGGCTGAGGTGGGCAGATT 83411

Qy   2058 GCTTGAGCACAGAGTTCCAGACACCAGCCTGAGCAACATGGCGAAATCCTGTCTCTTCAAG 2117
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  83412 GCTTGAGCACAGAGTTCCAGACACCAGCCTGAACAACATGGTGAAATCCTGTCTCTTCAAG 83471

Qy   2118 AAATAAAAATAATAATAATAAT 2138
          |||||  ||||| ||| ||| |
Db  83472 AAATATTTAAAAAAAAAAATT 83492
```

```
RESULT 4
BD183414
LOCUS       BD183414                  4807 bp    DNA       linear      PAT 17-JUN-2003
DEFINITION  Novel genes and proteins encoded by the genes.
ACCESSION   BD183414
VERSION     BD183414.1  GI:31875614
KEYWORDS    JP 200234592-A/127.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 4807)
  AUTHORS   Ohara,O., Nagase,T. and Nakajima,D.
  TITLE     Novel genes and proteins encoded by the genes
  JOURNAL   Patent: JP 200234592-A 127 03-DEC-2002;
            KAZUSA DNA RESEARCH INSTITUTE
COMMENT     OS   Homo sapiens (human)
            PN   JP 200234592-A/127
            PD   03-DEC-2002
            PF   26-FEB-2002 JP 2002049009
            PI   OSAMU OHARA,TAKAHIRO NAGASE,DAISUKE NAKAJIMA
            PC   C12N15/09,C07K14/47//A61K31/711,A61K38/00,A61K48/00,A61P25/00,
            PC   A61P25/14,
            PC   A61P25/18,A61P35/00,C12N15/00,A61K37/02
            CC   Novel genes and proteins encoded by the genes FH    Key
FEATURES             Location/Qualifiers
     source          1..4807
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9606"
     FT     CDS      (60)..(989).
            Location/Qualifiers

ORIGIN

Query Match            27.3%;  Score 588.6;  DB 6;  Length 4807;
Best Local Similarity   96.2%;  Pred. No. 4e-128;
Matches 603;  Conservative   0;  Mismatches  24;  Indels  0;  Gaps   0;

Qy    230 CCGGCAGCCGCAGCCGCCGCCGTCATGCCAGGCGCTGCTCGGCGAGCCAGAGGGATGG 289
```

```
Db        ||||   |||| |||| ||||   |||  |||||| |||  || .| ||   |     |||||||||||
      67  CCGGGGCCGCGCCTCCGCCCGCCGAAGCGCGGCGCCCACTGCCCAGAGCCAGAGGGATGG  126
Qy   290  TGGTAGTCACGGGCGGCGGGGAGCCAGACAGCCCGTCGTCAGGACGGTGCCATGTCCAGCTCTG  349
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   127  TGGTAGTCACGGGCGGCGGGGAGCCAGACAGCCCGTCGTCAGGACGGTGCCATGTCCAGCTCTG  186
Qy   350  ACGCCGAAGACGACTTTCTGGAGCCGCCACGCACGGCCACGGAGGCGGGGCACGCGC  409
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   187  ACGCCGAAGACGACTTTCTGGAGCCGCCACGCACGGCCACGGAGGCGGGGCACGCGC  246
Qy   410  TGCCCCTGCTGCCACAGGAGTTTCCTGAGGTTGTTCCCTTAACATCGGAGGGGCTCACT  469
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   247  TGCCCCTGCTGCCACAGGAGTTTCCTGAGGTTGTTCCCTTAACATCGGAGGGGCTCACT  306
Qy   470  TCACTACACGCTGTCCACACTGCGGTGCTACGAAGACACCATGTTGGCAGCCATGTTCA  529
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   307  TCACTACACGCTGTCCACACTGCGGTGCTACGAAGACACCATGTTGGCAGCCATGTTCA  366
Qy   530  GTGGGGCGGCACTACATCCCCACAGACTTCCGAGGGGCCGGTACTTCATCGACCGAGATGGCA  589
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   367  GTGGGGCGGCACTACATCCCCACAGACTTCCGAGGGGCCGGTACTTCATCGACCGAGATGGCA  426
Qy   590  CACACTTTGGAGATGTGCTAGATTTCCTGCTCAGGGGGACCTCCACCAGGGAGGCGTG  649
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   427  CACACTTTGGAGATGTGCTAGATTTCCTGCTCAGGGGGACCTCCACCAGGGAGGCGTG  486
Qy   650  TTCGAGCTGTGTACAAAGAGGCCCAGTACTATGCCATCGGGCCCCTCCTGGAGCAGCTGG  709
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   487  TTCGAGCTGTGTACAAAGAGGCCCAGTACTATGCCATCGGGCCCCTCCTGGAGCAGCTGG  546
Qy   710  AGAACATGCAGCCACTGAAGGGCGAGAAGGTGCGCCAAGCGTTTCTGGGACTCATGCCCT  769
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   547  AGAACATGCAGCCACTGAAGGGCGAGAAGGTGCGCCAAGCGTTTCTGGGACTCATGCCCT  606
Qy   770  ATTACAAAGACCACCACTTGGAGCCGATTGTGGAGATCGCCCGGCGTGCCGTCCAGCGGA  829
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   607  ATTACAAAGACCACCACTTGGAGCCGATTGTGGAGATCGCCCGGCGTGCCGTCCAGCGGA  666
Qy   830  AGGCCCCGCTTTGCCAAGCTCAAGAGCT  856
          ||||||||||||||||||||||||||||
Db   667  AGGCCCCGCTTTGCCAAGCTCAAGGTCT  693
```

RESULT 5
BC042482
LOCUS   BC042482   1211 bp   mRNA   linear   PRI 04-NOV-2003
DEFINITION Homo sapiens potassium channel tetramerisation domain containing 7,
            mRNA (cDNA clone MGC:34731 IMAGE:5165722), complete cds.
ACCESSION   BC042482
VERSION   BC042482.1 GI:27503737
KEYWORDS   MGC.
SOURCE   Homo sapiens (human)
ORGANISM   Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1 (bases 1 to 1211)
AUTHORS   Strausberg,R.L., Feingold,E.A., Grouse,L.H., Derge,J.G.,
            Klausner,R.D., Collins,F.S., Wagner,L., Shenmen,C.M., Schuler,G.D.,
            Altschul,S.F., Zeeberg,B., Buetow,K.H., Schaefer,C.F., Bhat,N.K.,
            Hopkins,R.F., Jordan,H., Moore,T., Max,S.I., Wang,J., Hsieh,F.,
            Diatchenko,L., Marusina,K., Farmer,A.A., Rubin,G.M., Hong,L.,
            Stapleton,M., Soares,M.B., Bonaldo,M.F., Casavant,T.L.,
            Scheetz,T.E., Brownstein,M.J., Usdin,T.B., Toshiyuki,S.,
            Carninci,P., Prange,C., Raha,S.S., Loquellano,N.A., Peters,G.J.,
            Abramson,R.D., Mullahy,S.J., Bosak,S.A., McEwan,P.J.,
            McKernan,K.J., Malek,J.A., Gunaratne,P.H., Richards,S.,
            Worley,K.C., Hale,S., Garcia,A.M., Gay,L.J., Hulyk,S.W.,
            Villalon,D.K., Muzny,D.M., Sodergren,E.J., Lu,X., Gibbs,R.A.,
            Fahey,J., Helton,E., Ketteman,M., Madan,A., Rodrigues,S.,
            Sanchez,A., Whiting,M., Madan,A., Young,A.C., Shevchenko,Y.,
            Bouffard,G.G., Blakesley,R.W., Touchman,J.W., Green,E.D.,
            Dickson,M.C., Rodriguez,A.C., Grimwood,J., Schmutz,J., Myers,R.M.,
            Butterfield,Y.S., Krzywinski,M.I., Skalska,U., Smailus,D.E.,
            Schnerch,A., Schein,J.E., Jones,S.J. and Marra,M.A.
TITLE   Generation and initial analysis of more than 15,000 full-length
            human and mouse cDNA sequences
JOURNAL   Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)
MEDLINE   22388257
PUBMED   12477932
REFERENCE   2 (bases 1 to 1211)
AUTHORS   Strausberg,R.
TITLE   Direct Submission
JOURNAL   Submitted (02-JAN-2003) National Institutes of Health, Mammalian
            Gene Collection (MGC), Cancer Genomics Office, National Cancer
            Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
            USA
REMARK   NIH-MGC Project URL: http://mgc.nci.nih.gov
COMMENT   Contact: MGC help desk
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: Life Technologies, Inc.
            cDNA Library Preparation: Life Technologies, Inc.
            cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
            DNA Sequencing by: Baylor College of Medicine Human Genome
            Sequencing Center
            Center code: BCM-HGSC
            Web site: http://www.hgsc.bcm.tmc.edu/cdna/
            Contact: amg@bcm.tmc.edu
            Gunaratne, P.H., Garcia, A.M., Lu, X., Hulyk, S.W., Loulseged, H.,
            Kowis, C.R., Sneed, A.J., Martin, R.G., Muzny, D.M., Nanavati,
            A.N., Gibbs, R.A.

            Clone distribution: MGC clone distribution information can be found
            through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov
            Series: IRAK Plate: 52 Row: 1 Column: 6
            This clone was selected for full length sequencing because it
            passed the following selection criteria: Hexamer frequency ORF
            analysis, GenomeScan gene prediction.
FEATURES   Location/Qualifiers
     source   1..1211
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="MGC:34731 IMAGE:5165722"
                     /tissue_type="Brain, adult medulla"
                     /clone_lib="NIH_MGC_119"
                     /lab_host="DH10B"
                     /note="Vector: pCMV-SPORT6"
     gene   1..1211
                     /gene="KCTD7"
                     /note="synonym: FLJ32069"
                     /db_xref="LocusID:154881"
     CDS   123..989
                     /codon_start=1
                     /product="KCTD7 protein"
                     /protein_id="AAH42482.1"
                     /db_xref="GI:27503738"
                     /db_xref="LocusID:154881"
                     /translation="MVVVTGREPDSRRQDGAMSSSDAEDDFLEPATPTATQAGHALPL
                     LPQEPFEVVPLNIGGAHFTTRLSTLRCYEDTMLAAMFSGRHYIPTDSEGRYFIDRRGT
                     HFGDVLNFLRSGDLPRERVRAVVKEAQYYAIGPLLEQLENMQPLKGEKVRQAFLGLM
                     PYYKDHLERIVEIARLRAVQRKARFAKLKVCVFKEMPITPYECPLLNSLRFERSESD
                     GQLFEHCEVDVSFGPWEAVADVDLLHCLVTDLSAQGLTVDHQCIGVCDKHLVNHYY
                     CKRPIYEFKITW"
     misc_feature   279..536
                     /note="K_tetra; Region: K+ channel tetramerisation domain.
                     The N-terminal, cytoplasmic tetramerisation domain (T1) of
                     voltage-gated K+ channels encodes molecular determinants
                     for subfamily-specific assembly of alpha-subunits into
                     functional tetrameric channels. It is distantly related to
                     the BTB/POZ domain pfam00651"
                     /db_xref="CDD:pfam0214"
ORIGIN

Query Match   27.3%;   Score 587; DB 9; Length 1211;
Best Local Similarity   96.0%;   Pred. No. 7.7e-128;

```
Matches   602;   Conservative   0;   Mismatches   25;   Indels   0;   Gaps   0;

Qy   230  CCGGCCAGCCCGCAGCCGCCGGCCGCGGTCATGCCAGGCGCTGCTCGCGGCGAGCCAGAGGGATGG  289
              ||||  |||  ||||  ||  |||  |||  ||  ||  |||  |  ||||||||||||||||
Db    67  CCGGGGCCGGCCCGCCCTCCGCCCGCCCGAAGCCGCCGCCACTGCCCCAGAGCCAGAGGGATGG  126

Qy   290  TGGTAGTCACGGGGCGGGGAGCCAGACAGCCGTCGTCAGGACGGTGCCATGTCCAGCTCTG  349
              ||||||||||||||||||||  ||||  ||  ||  |||||||||||||||||||||||
Db   127  TGGTAGTCACGGGGCGGGGAGCCAGACAGCCGTTGCTCAGGACGGTGCCATGTCCAGCTCTG  186

Qy   350  ACGCCGAAGACGACTTTCTGGAGCCGGGCCACGGCCACGGCGGGGCCACGGCCACGCGC  409
              ||||||||||||||||||||||||||||||||||||||||  ||  ||  ||  ||
Db   187  ACGCCGAAGACGACTTTCTGGAGCCGGGCCACGCCCAGGGGCCACGGCCACGCGC  246

Qy   410  TGCCCCTGCTGCCACAGGAGTTTCCTGAGTTGTTCCCCTTAAACATCGGAGGGGCTCACT  469
              |||||||||||||||||||||||  ||||||||||||||  |||  ||||  ||||
Db   247  TGCCCCTGCTGCCACAGGAGTTGCCTGAGTTGTTCCCCTTAAACATCGGAGGGGCTCACT  306

Qy   470  TCACTACACGCCTGTCCACACTGCGGTGCTACGAAGACACCATGTTGGCAGCCATGTTCA  529
              |||||||||||||||||||||||||||  |||  ||||||||||||||||||||||||
Db   307  TCACTACACGCCTGTCCACACTGCGGTGCTACGAAGACACCATGTTGGCAGCCATGTTCA  366

Qy   530  GTGGGCGGCACTACATCCCCACAGACTCCGAGGGCCGGTACTTCATCGACCGAGATGGCA  589
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   367  GTGGGCGGCACTACATCCCCACAGACTCCGAGGGCCGGTACTTCATCGACCGAGATGGCA  426

Qy   590  CACACTTTGGAGATGTGCTGAATTTCCTGCGCTCAGGGGACCTCCCACCCAGGGAGCGTG  649
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   427  CACACTTTGGAGATGTGCTGAATTTCCTGCGCTCAGGGGACCTCCCACCCAGGGAGCGTG  486

Qy   650  TTCGAGCTGTGTACAAAGAGGCCCAGTACTATGCCATCGGGCCCCTCCTGGAGCAGCTGG  709
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   487  TTCGAGCTGTGTACAAAGAGGCCCAGTACTATGCCATCGGGCCCCTCCTGGAGCAGCTGG  546

Qy   710  AGAACATGCAGCCACTGAAGGGCGAGAAGGTGCGCCAAGCGTTTCTGGGACTCATGCCCT  769
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   547  AGAACATGCAGCCACTGAAGGGCGAGAAGGTGCGCCAAGCGTTTCTGGGACTCATGCCCT  606

Qy   770  ATTACAAAGACCACTTGGAGCGGATTGTGGAGATCGCCCCGCTGCGTGCGGTCCAGCGGA  829
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   607  ATTACAAAGACCACTTGGAGCGGATTGTGGAGATCGCCCCGCTGCGTGCGGTCCAGCGGA  666

Qy   830  AGGCCCGCTTTGCCAAGCTCAAGAGCT  856
              |||||||||||||||  ||
Db   667  AGGCCCGCTTTGCCAAGCTCAAGGTCT  693


RESULT 6
AX714361
LOCUS       AX714361          2576 bp    DNA        linear    PAT 15-APR-2003
DEFINITION  Sequence 1045 from Patent EP1293569.
ACCESSION   AX714361
VERSION     AX714361.1  GI:29889313
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1
  AUTHORS   Isogai,T., Sugiyama,T., Otsuki,T., Wakamatsu,A., Sato,H., Ishii,S.,
            Yamamoto,J.I., Isono,Y., Hio,Y., Otsuka,K., Nagai,K., Irie,R.,
            Tamechika,I., Seki,N., Yoshikawa,T., Otsuka,M., Nagahari,K. and
            Masuho,Y.
  TITLE     Full-length cDNAs
  JOURNAL   Patent: EP 1293569-A 1045 19-MAR-2003;
            Helix Research Institute (JP) ; Research Association for
            Biotechnology (JP)
FEATURES             Location/Qualifiers
     source          1..2576
                     /organism="Homo sapiens"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:9606"
ORIGIN
```

```
Query Match                  27.3%;  Score 587;  DB 6;  Length 2576;
Best Local Similarity         96.0%;  Pred. No. 8.6e-128;
Matches  602;  Conservative  0;  Mismatches  25;  Indels  0;  Gaps  0;

Qy   230  CCGGCCAGCCCGCAGCCGCCGGCCGCGGTCATGCCAGGCGCTGCTCGCGGCGAGCCAGAGGGATGG  289
              ||||  |||  ||||  ||  |||  ||  ||  |||  |  ||||||||||||||||
Db    52  CCGGGGCCGGCCCGCCCTCCGCCCGCCCGAAGCCGCCGCCACTGCCCCAGAGCCAGAGGGATGG  111

Qy   290  TGGTAGTCACGGGGCGGGGAGCCAGACAGCCGTCGTCAGGACGGTGCCATGTCCAGCTCTG  349
              ||||||||||||||||||||  ||||  ||  ||  |||||||||||||||||||||||
Db   112  TGGTAGTCACGGGGCGGGGAGCCAGACAGCCGTTGCTCAGGACGGTGCCATGTCCAGCTCTG  171

Qy   350  ACGCCGAAGACGACTTTCTGGAGCCGGGCCACGGCCACGGCGGGGCCACGGCCACGCGC  409
              ||||||||||||||||||||||||||||||||||||||||  ||  ||  ||  ||
Db   172  ACGCCGAAGACGACTTTCTGGAGCCGGGCCACGCCCAGGGGCCACGGCCACGCGC  231

Qy   410  TGCCCCTGCTGCCACAGGAGTTTCCTGAGTTGTTCCCCTTAAACATCGGAGGGGCTCACT  469
              |||||||||||||||||||||||  ||||||||||||||  |||  ||||  ||||
Db   232  TGCCCCTGCTGCCACAGGAGTTGCCTGAGTTGTTCCCCTTAAACATCGGAGGGGCTCACT  291

Qy   470  TCACTACACGCCTGTCCACACTGCGGTGCTACGAAGACACCATGTTGGCAGCCATGTTCA  529
              |||||||||||||||||||||||||||  |||  ||||||||||||||||||||||||
Db   292  TCACTACACGCCTGTCCACACTGCGGTGCTACGAAGACACCATGTTGGCAGCCATGTTCA  351

Qy   530  GTGGGCGGCACTACATCCCCACAGACTCCGAGGGCCGGTACTTCATCGACCGAGATGGCA  589
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   352  GTGGGCGGCACTACATCCCCACAGACTCCGAGGGCCGGTACTTCATCGACCGAGATGGCA  411

Qy   590  CACACTTTGGAGATGTGCTGAATTTCCTGCGCTCAGGGGACCTCCCACCCAGGGAGCGTG  649
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   412  CACACTTTGGAGATGTGCTGAATTTCCTGCGCTCAGGGGACCTCCCACCCAGGGAGCGTG  471

Qy   650  TTCGAGCTGTGTACAAAGAGGCCCAGTACTATGCCATCGGGCCCCTCCTGGAGCAGCTGG  709
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   472  TTCGAGCTGTGTACAAAGAGGCCCAGTACTATGCCATCGGGCCCCTCCTGGAGCAGCTGG  531

Qy   710  AGAACATGCAGCCACTGAAGGGCGAGAAGGTGCGCCAAGCGTTTCTGGGACTCATGCCCT  769
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   532  AGAACATGCAGCCACTGAAGGGCGAGAAGGTGCGCCAAGCGTTTCTGGGACTCATGCCCT  591

Qy   770  ATTACAAAGACCACTTGGAGCGGATTGTGGAGATCGCCCCGCTGCGTGCGGTCCAGCGGA  829
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   592  ATTACAAAGACCACTTGGAGCGGATTGTGGAGATCGCCCCGCTGCGTGCGGTCCAGCGGA  651

Qy   830  AGGCCCGCTTTGCCAAGCTCAAGAGCT  856
              |||||||||||||||  ||
Db   652  AGGCCCGCTTTGCCAAGCTCAAGGTCT  678


RESULT 7
AK056631
LOCUS       AK056631          2576 bp    mRNA       linear    PRI 01-AUG-2002
DEFINITION  Homo sapiens cDNA FLJ32069 fis, clone OCBBF1000118, weakly similar
            to TUMOR NECROSIS FACTOR, ALPHA-INDUCED PROTEIN 1, ENDOTHELIAL.
ACCESSION   AK056631
VERSION     AK056631.1  GI:16552086
KEYWORDS    oligo capping; fis (full insert sequence).
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1
  AUTHORS   Oshima,A., Takahashi-Fujii,A., Tanase,T., Imose,N., Takeuchi,K.,
            Arita,M., Musashino,K., Yuuki,H., Hara,H., Sugiyama,T., Irie,R.,
            Otsuki,T., Sato,H., Ota,T., Wakamatsu,A., Ishii,S., Yamamoto,J.,
            Isono,Y., Kawai-Hio,Y., Saito,K., Nishikawa,T., Kimura,K.,
            Yamashita,H., Matsuo,K., Nakamura,Y., Sekine,M., Kikuchi,H.,
            Kanda,K., Wagatsuma,M., Murakawa,K., Kanehori,K., Sugiyama,A.,
            Kawakani,B., Suzuki,S., Nagahari,K., Masuho,Y., Nagai,K.
            and Isogai,T.
  TITLE     NEDO human cDNA sequencing project
  JOURNAL   Unpublished
```

REFERENCE   2 (bases 1 to 2576)
AUTHORS   Isogai,T., Otsuki,T. and Sugiyama,T.
TITLE   Direct Submission
JOURNAL   Submitted (24-OCT-2001) Takao Isogai, Helix Research Institute,
            Genomics Laboratory; 1532-3 Yana, Kisarazu, Chiba 292-0812, Japan
            (E-mail:genomics@hri.co.jp, Tel:81-438-52-3975, Fax:81-438-52-3986)
COMMENT   NEDO human cDNA sequencing project supported by Ministry of
            Economy, Trade and Industry of Japan; cDNA full insert sequencing:
            Research Association for Biotechnology (RAB); cDNA full insert sequencing:
            construction: Helix Research Institute (HRI) (supported by Japan
            Key Technology Center etc.); 5'- & 3'-end one pass sequencing: RAB,
            HRI, and Biotechnology Center, National Institute of Technology and
            Evaluation; clone selection for full insert sequencing: RAB and
            HRI.

FEATURES   Location/Qualifiers
    source   1..2576
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="OCBBF1000118"
                     /tissue_type="brain"
                     /clone_lib="OCBBF1"
                     /dev_stage="fetus"
                     /note="cloning vector: pME18SFL3"
    CDS   108..977
                     /note="unnamed protein product"
                     /codon_start=1
                     /protein_id="BAB71236.1"
                     /db_xref="GI:16552087"
                     /translation="MVVVTGREPDSRRQDGAMSSDAEDDFLEPATPIATQAGHALPL
                     LPQEFPEVVPLNIGAHFTTRLSTLRCYEDTMLAAMFSGRHYIPTDSEGRYFIDRDGT
                     HFGDVLNFLRSGDLPPRERVRAVVKEAQYYAIGPLLEQLENMQPLKGEKVRQAFLGLM
                     PYYKDHLERIVEIARLRAVORKARFAKLKVCVFKEEMPITPYECPLLNSLRFERSESD
                     GQLFEHCEVDVSFGPWEAVADVYDLLHCIVTDLSAQGLTVDHQCIGVCDKHLVNHYY
                     CKRPIYEFKITWW"

ORIGIN

Query Match   27.3%; Score 587; DB 9; Length 2576;
Best Local Similarity   96.0%; Pred. No. 8.6e-128;
Matches 602; Conservative   0; Mismatches   25; Indels   0; Gaps   0;

Qy   230 CCGGCCAGCCCAGCCGGCCCGCCGTCATGCCAGGCCTGCTCGGCAGCGGCCAGAGGGATGG 289
          ||||| ||||||||| ||||||| | | | | | |||||||||||||||||||
Db   52 CCGGGGCCGCGCCGCTCCGCCCCGAGCCCCCACTGCCCAGCGCCCCAGAGCCAGAGGGATGG 111

Qy   290 TGGTAGTCACGGGGCGGGGAGCCCAGACAGCCGTGTCAGGACGGTGCCATGTCCAGCTCTG 349
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   112 TGGTAGTCACGGGGCGGGGAGCCCAGACAGCCGTGTCAGGACGGTGCCATGTCCAGCTCTG 171

Qy   350 ACGCCGAAGACAGACTTTCTGGAGCCGGACTCCGAGGCGCACGCAGCGGGGCCACGCGC 409
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   172 ACGCCGAAGACAGACTTTCTGGAGCCGGACTCCGAGGCGCACGCAGCGGGGCCACGCGC 231

Qy   410 TGCCCCTGCTGCCACAGAGTTTCCTGAGGTTGTTCCCTTAACATCGGAGGGGCTCACT 469
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   232 TGCCCCTGCTGCCACAGAGTTTCCTGAGGTTGTTCCCTTAACATCGGAGGGGCTCACT 291

Qy   470 TCACTACACGCCTGTCCACACTGCGGTGCTACGAAGACACCACCATGTTGGCAGCCATGTTCA 529
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   292 TCACTACACGCCTGTCCACACTGCGGTGCTACGAAGACACCACCATGTTGGCAGCCATGTTCA 351

Qy   530 GTGGGGCGGCACTACATCCCACAGACTCCGAGGCGTACTTCATCGACCGAGATGCA 589
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   352 GTGGGGCGGCACTACATCCCACAGACTCCGAGGCGTACTTCATCGACCGAGATGCA 411

Qy   590 CACACTTTGGAGATGTGCTGAATTCCTGCGCTCAGGGAGACCTCCACCCACCAGGGAGCGTG 649
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   412 CACACTTTGGAGATGTGCTGAATTCCTGCGCTCAGGGAGACCTCCACCCACCAGGGAGCGTG 471

Qy   650 TTCGAGCTGTGTACAAAGAGGCCCAGTACTATGCCATGGGCATCGGGCCCCTGCTGGAGCAGCTGG 709
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   472 TTCGAGCTGTGTACAAAGAGGCCCAGTACTATGCCATGGGCATCGGGCCCCTGCTGGAGCAGCTGG 531

Qy   710 AGAACATGCAGCCACTGAAGGGCGAGAAGGTGCGCCAAGCGTTCTTGGGACTCATGCCCT 769
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   532 AGAACATGCAGCCACTGAAGGGCGAGAAGGTGCGCCAAGCGTTCTTGGGACTCATGCCCT 591

Qy   770 ATTACAAAGACCACTTGGAGCGGATTGTGGAGATCGCCCGGCTGCGCGTCGGTCCAGCGGA 829
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   592 ATTACAAAGACCACTTGGAGCGGATTGTGGAGATCGCCCGGCTGCGCGTCGGTCCAGCGGA 651

Qy   830 AGGCCCGCTTTGCCAAGCTCAAGAGCT 856
          ||||||||||||||||||||||| ||
Db   652 AGGCCCGCTTTGCCAAGCTCAAGGTCT 678

RESULT 8
BD275557
LOCUS   BD275557   1124 bp   DNA   linear   PAT 17-JUL-2003
DEFINITION MOLECULES OF THE IMMUNE SYSTEM.
ACCESSION   BD275557
VERSION   BD275557.1 GI:33085325
KEYWORDS   JP 2002540791-A/2.
SOURCE   Homo sapiens (human)
ORGANISM   Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1 (bases 1 to 1124)
AUTHORS   Lu,D.A.M., Azimzai,Y., Baughn,M.R., Tang,T.Y., Lal,P. and Yue,H.
TITLE   MOLECULES OF THE IMMUNE SYSTEM
JOURNAL   Patent: JP 2002540791-A 2 03-DEC-2002;
            INCYTE PHARMACEUTICALS INC.Henry YUE,Preeti LAL,Tom Y TANG, Mariah
            R BAUGHN,Yalda AZIMZAI,Dyung Aina M LU
COMMENT   OS   Homo sapiens
            PN   JP 2002540791-A/2
            PD   03-DEC-2002
            PF   04-APR-2000 JP 2000609571
            PR   05-MAY-1999 US   60/132647,05-APR-1999 US   60/127852 PI
            dyung aina m lu,yalda azimzai,mariah r baughn,tom y tang, PI
            preeti lal, henry yue
            CC   This description about <220> can't be interpreted CC   <220>
            CC   <221> misc_feature
            CC   <223> Incyte ID No.: 2751129CB1
            FH   Key   Location/Qualifiers
FEATURES   Location/Qualifiers
    source   1..1124
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9606"
ORIGIN

Query Match   26.2%; Score 565.2; DB 6; Length 1124;
Best Local Similarity   99.3%; Pred. No. 1.1e-122;
Matches 578; Conservative   0; Mismatches   3; Indels   1; Gaps   1;

Qy   275 GAGCCAGAGGGATGGTGGTAGTCACGGGGCGGGGAGCCCAGACAGCCGTGTCAGGACGGTG 334
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   17 GAGCCAGAGGGATGGTGGTAGTCACGGGGCGGGGAGCCCAGACAGCCGTGTCAGGACGGTG 76

Qy   335 CCATGTCCAGCTCTGACGCCGAAGACAGACTTTCTGGAGCCGGACTTTCTGGAGGCGCCACGC 394
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   77 CCATGTCCAGCTCTGACGCCGAAGACAGACTTTCTGGAGCCGGACTTTCTGGAGGCGCCACGC 136

Qy   395 AGGCGGGGCCACGCGCTGCCCCTGCTGCCACAGAGTTCCTGAGGTTGTTCCCTTAACA 454
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   137 AGGCGGGGCCACGCGCTGCCCCTGCTGCCACAGAGTTCCTGAGGTTGTTCCCTTAACA 196

Qy   455 TCGGAGGGGCTCACTTCACTACACGCCTGTCCACACTGCGGTGCTACGAAGACACCATGT 514
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   197 TCGGAGGGGCTCACTTCACTACACGCCTGTCCACACTGCGGTGCTACGAAGACACCATGT 256

Qy   515 TGGCAGCCATGTTCAGTGGGCGGCACTACATCCCACAGACTCCGAGGCCGGTACTTCA 574
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   257 TGGCAGCCATGTTCAGTGGGCGGCACTACATCCCACAGACTCCGAGGCCGGTACTTCA 316

```
Qy    575  TCGACCGAGATGGCACACACTTTGGAGATGTGCTGAATTTCCTGCGCCTCAGGGGACCTCC  634
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    317  TCGACCGAGATGGCACACACTTTGGAGATGTGCTGAATTTCCTGCGCCTCAGGGGACCTCC  376

Qy    635  CACCCAGGAGCGGCGTGTTCGAGCTGTGTACAAAGAGGCCCAGTACTACACGCCTGGGCCC  694
                |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    377  CACCCAGGAGCGGCGTGTTCGAGCTGTGTACAAAGAGGCCCAGTACTACACGCCTGGGCCC  436

Qy    695  TCCTGGAGCAGCTGGAGAACATGCAGCCACTGAAGGGCGAGAAGGTGCGCCAAGCGTTTC  754
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    437  TCCTGGAGCAGCTGGAGAACATGCAGCCACTGAAGGGCGAGAAGGTGCGCCAAGCGTTTC  496

Qy    755  TGGGACTCATGCCCTATTACAAAGACCACTTGGAGCGGATTGTGGAGATCGCCCGGCTGC  814
                |||||||||||||||||||||||||||||||||||||||||||||||||||||||||| |
Db    497  TGGGACTCATGCCCTATTACAAAGACCACTTGGAGCGGATTGTGGAGATCG--CGGGCTGC  555

Qy    815  GTGCGGTCCAGCGGAAGGCCCGCTTGCCAAGCTCAAGAGCT  856
                ||||||||||||||||||||||||||||||||| |||||
Db    556  GTGCGGTCCAGCGGAAGGCCCGCTTGCCAAGCTCAAGGTCT  597
```

RESULT 9
AB056802

LOCUS   AB056802   3864 bp   mRNA   linear   PRI 10-JAN-2002
DEFINITION   Macaca fascicularis brain cDNA clone:QflA-12743, full insert
              sequence.
ACCESSION   AB056802
VERSION   AB056802.1 GI:13365904
KEYWORDS   fis (full insert sequence); oligo capping.
SOURCE   Macaca fascicularis (crab-eating macaque)
ORGANISM   Macaca fascicularis
              Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
              Mammalia; Eutheria; Primates; Catarrhini; Cercopithecidae;
              Cercopithecinae; Macaca.
REFERENCE   1
AUTHORS   Osada,N., Hida,M., Kusuda,J., Tanuma,R., Hirata,M., Hirai,M.,
              Terao,K., Suzuki,Y., Sugano,S. and Hashimoto,K.
TITLE   Prediction of unidentified human genes on the basis of sequence
              similarity to novel cDNAs from cynomolgus monkey brain
JOURNAL   Online Publication
REMARK   Genome Biology 2001 3(1); research0006.1-0006.5; *
              http://genomebiology.com/2001/3/1/research/0006/
REFERENCE   2 (bases 1 to 3864)
AUTHORS   Hashimoto,K., Osada,N., Hida,M., Kusuda,J. and Sugano,S.
TITLE   Direct Submission
JOURNAL   Submitted (09-MAR-2001) Katsuyuki Hashimoto, National Institute of
              Infectious Diseases, Division of Genetic Resources; 23-1, Toyama
              1-chome, Shinjuku-ku, Tokyo 162-8640, Japan
              (E-mail:khashi@nih.go.jp; URL:http://www.nih.go.jp/yoken/genebank/,
              Tel:81-3-5285-1111(ex.2120), Fax:81-3-5285-1181)
              Lab host: TOP10
              Vector: pME18S-FL3 (Acc.No. AB009864)
              R. Site1:   DraIII (CACTGTGTG)
              R. Site2:   DraIII (CACCATGTG)
              Description: 1st strand cDNA was primed with an oligo(dT) primer
              [ATGTGGCCCTTTTTTTTTTTTTTTT]; double-stranded cDNA was synthesized
              using specific 5' and 3' primers and amplified by PCR. The PCR
              product was digested with SfiI and size selection was performed to
              exclude fragments <1.5kb.The SfiI-digested PCR product was cloned
              into distinct DraIII sites of pME18S-FL3. XhoI sites just outside
              the DraIII sites can be used to isolate the cDNA insert. Libraries
              were constructed by Sugano et al.[University of Tokyo, Institute of
              Medical Science]. Custom primer used for sequencing
              (5' end primer [CTTCTGCCTCTAAAAGCTGCG];
              3' end primer [CGACCTGCAGCTCGAGCACA]).

COMMENT

FEATURES   Location/Qualifiers
source   1..3864
              /organism="Macaca fascicularis"
              /mol_type="mRNA"
              /db_xref="taxon:9541"
              /clone="QflA-12743"
              /sex="male"
              /tissue_type="frontal lobe left"

              /clone_lib="macaque brain cDNA library QflA"
              /dev_stage="adult"
              /note="Host: TOP10; Vector: pME18S-FL3 (Acc.No. AB009864);
              R. Site1: DraIII (CACTGTGTG); R. Site2: DraIII
              (CACCATGTG)"
CDS   167..811
              /codon_start=1
              /product="hypothetical protein"
              /protein_id="BAB39326.1"
              /db_xref="GI:13365905"
              /translation="MLAAMFSGRHYIPTDSEGRYLIDRDGTHFGDVLNFLRSGDLPPR
              ERVRAVYKEAQYAIGPLLEQLENMQPLKGEKVRQAFLGLMPYYKDHLERIVEIARLR
              AVQRKARFAKLKVCVFKEEMPITPYEQLLNSLRFERSESDGQLFEHHCEVDVSFGPW
              EAVADVDLLHCLVTDLSAQGLTVDHQCIGVCDKHLVNHYYCKRPIYEFKITWW"

ORIGIN

Query Match   20.1%; Score 433.6; DB 9; Length 3864;
Best Local Similarity   93.0%; Pred. No. 1.6e-91;
Matches 454; Conservative   0; Mismatches 34; Indels   0; Gaps   0;

```
Qy    369  GGAGCGGCCGCCACGCCGACGGCCAGCGAGGGGGGGCCGCTGCCCCTGCTGCCACAGGA  428
                 ||||          |||      ||||||||||||||||| || ||||||||||||
Db     25  GGAGTGCCCCGGGCCGCACCGCTCCGCCTCCGGCCAGCCGCCCCACTGACCGCCCCAGGA   84

Qy    429  GTTCCTGAGGTTGTTCCCCTTAACATCGGAGGGGCTCACTTCACTACACGCCTGTCCAC  488
                 |||||||||||||||||| ||||||||||||||||||||||| ||||||||||||||
Db     85  GTTCCTGAGGTTGTTCCGCTTAACATCGGAGGGGCTCACTTTACTACACGCCTGTCCAC  144

Qy    489  ACTGCGGGTGCTACGAAGACACCATGTTGGCAGCCATGTTCAGTGGCGGGCGACTACATCCC  548
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    145  ACTGCGGGTGCTACGAAGACACCATGTTGGCAGCCATGTTCAGTGGCGGGCGACTACATCCC  204

Qy    549  CACAGACTTCGAGGGCCGTGACTTCATCGACCGAGATGGCACCACTTTTGGAGATGTGCT  608
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    205  CACGGACTTCGAGGGCCGTGACTTCATCGACCGAGATGGCACCACACTTTTGGAGATGTGCT  264

Qy    609  GAATTTCCTGCGCTCAGGGGACCTCCCACCCAGGAGCGGTGTTCGAGCTGTGTACAAAGA  668
                 |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    265  GAATTTCCTGCGCTCAGGGGACCTCCCACCCCAGGAGCGGTGTTCGAGCTGTGTACAAAGA  324

Qy    669  GGCCCAGTACTACTACATGCCATCGGGCCCTCGGCGCCCTCCTGGAGCAGCTGGAGAACATGCAGCCACTGAA  728
                 |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    325  GGCCCAGTACTACTACATGCCATCGGGCCCTCGGCGCCCTCCTGGAGCAGCTGGAGAACATGCAGCCACTGAA  384

Qy    729  GGGCGAGAAGGTGCGCCAAGCGTTTCTGGGACTCATGCCCTATTACAAAGACCACTTGGA  788
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    385  GGGCGAGAAGGTGCGCCAAGCGTTTCTGGGACTCATGCCCTATTACAAAGACCACTTGGA  444

Qy    789  GCGGATTGTGGAGATCGCCCGGCTGCGTGCGGTCCAGCGGAAGGCCCGCTTGCCAAGCT  848
                 |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    445  GCGGATTGTGGAGATCGCCCGGCTGCGTGCGGTCCAGCGGAAGGCCCGCTTGCCAAGCT  504

Qy    849  CAAGAGCT  856
                 |||| |||
Db    505  CAAGGTCT  512
```

RESULT 10
AX575790

LOCUS   AX575790   680 bp   DNA   linear   PAT 07-JAN-2003
DEFINITION   Sequence 8 from Patent WO02068626.
ACCESSION   AX575790
VERSION   AX575790.1 GI:27552276
KEYWORDS   .
SOURCE   Homo sapiens (human)
ORGANISM   Homo sapiens
              Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
              Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1
AUTHORS   Friddle,C.J., Gerhardt,B., Hilbun,E. and Turner,C.A.
TITLE   Novel human ion channel-related proteins and polynucleotides
              encoding the same
JOURNAL   Patent: WO 02068626-A 8 06-SEP-2002;

Primer B: GAGTATTCCAACACTTCATGGGC
STS size: 325
PCR Profile:
Initial incubation: 95 degrees C for 10 minutes
Denaturation: 94 degrees C for 30 seconds
Annealing: 60 degrees C for 30 seconds
Polymerization: 72 degrees C for 23 seconds
PCR Cycles: 30
Thermal Cycler: Perkin Elmer 9700

Protocol:
Template: 25 ng
Primer: each 1 uM
dNTPs: each 200 uM
AmpliTaq Gold Polymerase: 0.07 units/ul
Total Vol: 5 ul

Buffer:
MgCl2: 2.5 mM
KCl: 50 mM
Tris-HCl: 10 mM
pH: 8.3

BAC ends sequenced at TIGR from the RPCI11 BAC library. Designed
and developed at the Stanford Human Genome Center.

FEATURES   Location/Qualifiers
source   1..374
                  /organism="Homo sapiens"
                  /mol_type="genomic DNA"
                  /db_xref="taxon:9606"
                  /map="7"
                  /clone_lib="Human"
STS   20..344
primer_bind   20..42
primer_bind   complement(322..344)
ORIGIN

Query Match   17.4%;   Score 374;   DB 11;   Length 374;
Best Local Similarity   100.0%;   Pred. No. 1.4e-77;
Matches 374;   Conservative   0;   Mismatches   0;   Indels   0;   Gaps   0;

Qy   977 AGAAGAAATCCCCTGTCCAGTTGCCTGCAGGAGTATTCCAACACTTCATGGGCTAGAGGA 1036
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   374 AGAAGAAATCCCCTGTCCAGTTGCCTGCAGGAGTATTCCAACACTTCATGGGCTAGAGGA 315

Qy   1037 TTCCATTGAGATGGGGTTTACGTCTTGATTTTGAACACCTGTCAGCACTGTTCTCTGTTT 1096
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   314 TTCCATTGAGATGGGGTTTACGTCTTGATTTTGAACACCTGTCAGCACTGTTCTCTGTTT 255

Qy   1097 GCATGGCAATTCTGACCCTTTTATGCAACAACACCCCTGGGACAACACCCCAGATTTGTAGA 1156
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   254 GCATGGCAATTCTGACCCTTTTATGGCAACAACACCCCTGGGACAACACCCCAGATTTGTAGA 195

Qy   1157 TTGAGATCCAAAGGTAGAAATTCCAGACAGTCCAACAAGTATCAAGTGATGTGTTTCCAG 1216
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   194 TTGAGATCCAAAGGTAGAAATTCCAGACAGTCCAACAAGTATCAAGTGATGTGTTTCCAG 135

Qy   1217 AGTGGAAGGCTCTCACCGTGTCCCAGATTTCTGGGGTTTGTAAGCAGTACTGGCCATTT 1276
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   134 AGTGGAAGGCTCTCACCGTGTCCCAGATTTCTGGGGTTTGTAAGCAGTACTGGCCATTT 75

Qy   1277 GTGACCCTGTTTTTACCTAATCATTCTGTCTTTTTAGGACATGGTTTTACCCGATCCCT 1336
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   74 GTGACCCCTGTTTTTTACCTAATCATTCTGTCTTTTTTAGGACATGGTTTTACCCGATCCCT 15

Qy   1337 GGCAAAGGATCAG 1350
          |||||||||||||
Db   14 GGCAAAGGATCCAG 1

RESULT 12
AX575786
LOCUS   AX575786   363 bp   DNA   linear   PAT 07-JAN-2003
DEFINITION Sequence 4 from Patent WO02068626.

FEATURES   Lexicon Genetics Incorporated (US)
source   Location/Qualifiers
                  1..680
                  /organism="Homo sapiens"
                  /mol_type="unassigned DNA"
                  /db_xref="taxon:9606"
ORIGIN

Query Match   18.0%;   Score 387.8; DB 6;   Length 680;
Best Local Similarity   88.2%;   Pred. No. 8.2e-81;
Matches 456;   Conservative   1;   Mismatches   3;   Indels   57;   Gaps   1;

Qy   146 CGGGTCAGGCCCCAGCTGGGCGGGCGGAGCGGGTCGGCGTTGAGGGAGGCCACCGCCCTCCCGC 205
          ||||||||||||||||||||||||||||||||||| |||||||||||||||||||||||
Db   1 CGGGTCAGGCCCCAGCTGGGCGGGCGGAGCGGGTCGGCGTTGAGGGAGGCCACCGCCCTCCCGC 60

Qy   206 CTGCGCACTGCCTCTCGCCCCCTCCGGCCAGCCCCCAGCGCCGGCCGCGTCATGCCAGGGCG 265
          |||||||||||||||||||||||||||||||||||||||| : |||||||||||||||
Db   61 CTGCGCACTGCCTCTCGCCCCCTCCGGCCAGCCCCCAGCGCCGGCCGCGTSATGCCAGGGCG 120

Qy   266 CTGCTCGGC-- 274
          |||||||||
Db   121 CTGCTCGGCGTGAGGGAGTGCCCGGGGCCGCCGYCTCCGCCCGCCCGGAAGCCGCGGCCCAC 180

Qy   275 -----GAGCCAGGAGGGATGGTGGTAGTCACGGGGCGGGAGCCGAGCCAGGGCGTCAGG 328
               |||||||||||||||||||||||||||||||||||||||||||||||||
Db   181 TGCCCAGAGCCAGGAGGGATGGTGGTAGTCACGGGGCGGGAGCCAGACAGCCGTCGTCAGG 240

Qy   329 ACGGTGCCATGTCCAGCTCTGACGCCGAAGACGACTTTCTGGAGCCGACGGCCACCGACGG 388
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   241 ACGGTGCCATGTCCAGCTCTGACGCCGAAGACGACTTTCTGGAGCCGACGGCCACCGACGG 300

Qy   389 CCACGGCAGGCGGGGCACGCGCGTGCCCCTGCTGCCACAGGAGTTTCCTGAGGTTGTGTTCCCC 448
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   301 CCACGGCAGGCGGGGCACGCGCGTGCCCCTGCTGCCACAGGAGTTTCCTGAGGTTGTGTTCCCC 360

Qy   449 TTAACATCGGAGGGGCTCACTTCACTACACGGCCTGCCACACGCCGTGCTACGAAGACA 508
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   361 TTAACATCGGAGGGGCTCACTTCACTACACGCCTGCCACACGCCGTGCTACGAAGACA 420

Qy   509 CCATGTTGGCAGCCATGTTCAGTGGCGGCACTACACATCCCCACACAGACTCCGAGGGCCGGT 568
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   421 CCATGTTGGCAGCCATGTTCAGTGGCGGCACTACACATCCCCACACAGACTCCGAGGGCCGGT 480

Qy   569 ACTTCATCGACCGAGATGGCACACACTTTGGAGATGT 605
          |||||||||||||||||||||||||||||||||||||
Db   481 ACTTCATCGACCGAGATGGCACACACTTTGGGTATGT 517

RESULT 11
G51111/c
LOCUS   G51111   374 bp   DNA   linear   STS 30-MAR-2000
DEFINITION SHGC-80071 Human Homo sapiens STS genomic, sequence tagged site.
ACCESSION   G51111
VERSION   G51111.1 GI:5222288
KEYWORDS   STS.
SOURCE   Homo sapiens (human)
ORGANISM   Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1 (bases 1 to 374)
AUTHORS   Olivier,M. and Cox,D.R.
TITLE   Unpublished (2000)
JOURNAL   Unpublished

COMMENT   Contact: Michael Olivier, David R. Cox
            Stanford Human Genome Center
            Stanford University School of Medicine
            4005 Miranda Ave. 2nd Fl., Palo Alto, CA 94025, USA
            Tel: (650) 320-5800
            Fax: (650) 320-5801
            Email: olivier@shgc.stanford.edu
            Primer A: TCGGGTAAAACCATGTCCTAAAA

```
REFERENCE    4  (bases 1 to 188791)
AUTHORS      McPherson,J.D. and Waterston,R.H.
TITLE        Direct Submission
JOURNAL      Submitted (21-JUN-2002) Genome Sequencing Center, 4444 Forest Park
             Parkway, St. Louis, MO 63108, USA
REFERENCE    5  (bases 1 to 188791)
AUTHORS      Wilson,R.
TITLE        Direct Submission
JOURNAL      Submitted (05-NOV-2003) Department of Genetics, Washington
             University, 4444 Forest Park Avenue, St. Louis, Missouri 63108, USA
             On Jun 21, 2002 this sequence version replaced gi:2006729.
COMMENT      ---------- Genome Center
             Center: Washington University Genome Sequencing Center
             Center code: WUGSC
             Web site: http://genome.wustl.edu
             Contact: submissions@watson.wustl.edu
             ---------- Summary Statistics
             Center project name: M_BA0037M07
             ----------

NOTICE:      This sequence may not represent the entire insert of this
             clone. It may be shorter because we only sequence overlapping
             clone sections once, or longer because we provide a small overlap
             between neighboring data submissions.

             This sequence was finished as follows unless otherwise noted:
             all regions were double stranded, sequenced with an alternate
             chemistry, or covered by high quality data (i.e., phred quality >=
             30); an attempt was made to resolve all sequencing problems, such
             as compressions and repeats; all regions were covered by sequence
             from more than one subclone; and the assembly was confirmed by
             restriction digest.

MAPPING INFORMATION:
             Mapping information for this clone was provided by Dr. Wes Warren,
             Department of Genetics, Washington University, St. Louis MO. For
             additional information about the map position of this sequence, see
             http://genome.wustl.edu

SOURCE INFORMATION:
             The RPCI-23 BAC Library has been constructed by Kazutoyo Osegawa
             and Minako Tateno in the laboratory of Pieter de Jong
             (http://www.chori.org) from female C57BL/6J mouse kidney and/or
             brain genomic DNA. The clone and detailed information can be
             obtained from Research Genetics, Inc. (http://www.resgen.com) or
             Pieter de Jong and coworkers at http://www.chori.org

NEIGHBORING SEQUENCE INFORMATION:
             This sequence is the entire insert of the clone. This clone is
             overlapped by AC122830.
FEATURES             Location/Qualifiers
     source          1..188791
                     /organism="Mus musculus"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:10090"
                     /chromosome="8"
                     /map="8"
                     /clone="RP23-37M7"
                     /clone_lib="RPCI-23"
     repeat_region   6..531
                     /rpt_family="L1"
     repeat_region   744..1012
                     /rpt_family="MaLR"
     repeat_region   1386..2293
                     /rpt_family="L1"
     repeat_region   2343..2550
                     /rpt_family="B2"
     repeat_region   3654..3800
                     /rpt_family="B4"
     repeat_region   3835..3997
                     /rpt_family="Alu"
     repeat_region   4486..4563
                     /rpt_family="ERV1"
```

```
ACCESSION   AX575786
VERSION     AX575786.1  GI:27552274
KEYWORDS    .
SOURCE      Homo sapiens (human)
ORGANISM    Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1
AUTHORS     Friddle,C.J., Gerhardt,B., Hilbun,E. and Turner,C.A.
TITLE       Novel human ion channel-related proteins and polynucleotides
            encoding the same
JOURNAL     Patent: WO 02068626-A 4 06-SEP-2002;
            Lexicon Genetics Incorporated (US)
FEATURES             Location/Qualifiers
     source          1..363
                     /organism="Homo sapiens"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:9606"
ORIGIN

Query Match        14.6%;  Score 315.2; DB 6;  Length 363;
Best Local Similarity  99.1%;  Pred. No. 1.1e-63;
Matches   317;  Conservative  0;  Mismatches  3;  Indels  0;  Gaps   0;

Qy    286 ATGGTGGTAGTCACGGGGCGGGGAGCCAGACAGCCGTCGTCAGGACGGTGCCATGTCCAGC 345
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1 ATGGTGGTAGTCACGGGGCGGGGAGCCAGACAGCCGTCGTCAGGACGGTGCCATGTCCAGC 60

Qy    346 TCTGACGCCGAAGACGACTTTCTGGAGCCGGCCACGCCACGCAGGCGGGGGCAC 405
              |||||||||||||||||||||||||||||| |||||||||||||||| |||||||
Db     61 TCTGACGCCGAAGACGACTTTCTGGAGCCGCCCACGCCACGCCAGCGCGGGGCAC 120

Qy    406 GCGCTGCCCCTGCTGCCACGAGAGTTTCCTGAGGTTGTTCCCCTTAACATCGGAGGGGCT 465
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    121 GCGCTGCCCCTGCTGCCACAGAGTTTCCTGAGGTTGTTCCCCTTAACATCGGAGGGGCT 180

Qy    466 CACTTCACTCACACGCCTGTCCACACTGCGGTGCTACGAAGACACCATGTTGGCAGCCATG 525
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    181 CACTTCACTCACACGCCTGTCCACACTGCGGTGCTACGAAGACACCATGTTGGCAGCCATG 240

Qy    526 TTCAGTGGCGGCACTACTACCCCACAGACTCCGAGGCCGTACTTCATCGACCGAGAT 585
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    241 TTCAGTGGGCGGCACTACATCCCCACGGACTCCGAGGCCGGTACTTCATCGACCGAGAT 300

Qy    586 GGCACACACTTTGGAGATGT 605
              ||||||||||||||||||||
Db    301 GGCACACACTTTGGGTATGT 320
```

```
RESULT 13
AC117185/c
LOCUS       AC117185            188791 bp    DNA     linear   ROD 05-NOV-2003
DEFINITION  Mus musculus BAC clone RP23-37M7 from 8, complete sequence.
ACCESSION   AC117185
VERSION     AC117185.2  GI:21536193
KEYWORDS    HTG.
SOURCE      Mus musculus (house mouse)
ORGANISM    Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 188791)
AUTHORS     Haglund,K. and Schatzkamer,K.
TITLE       The sequence of Mus musculus BAC clone RP23-37M7
JOURNAL     Unpublished (2001)
REFERENCE   2  (bases 1 to 188791)
AUTHORS     Wilson,R.
TITLE       Sequencing of Mus musculus
JOURNAL     Unpublished (2001)
REFERENCE   3  (bases 1 to 188791)
AUTHORS     McPherson,J.D. and Waterston,R.H.
TITLE       Direct Submission
JOURNAL     Submitted (08-APR-2002) Genome Sequencing Center, 4444 Forest Park
            Parkway, St. Louis, MO 63108, USA
```

```
repeat_region   4936..5021
                /rpt_family="Alu"
repeat_region   4955..5024
                /rpt_family="B4"
repeat_region   7704..7831
                /rpt_family="B4"
repeat_region   8033..8179
                /rpt_family="Alu"
repeat_region   9503..9572
                /rpt_family="Alu"
repeat_region   10671..10781
                /rpt_family="ID"
repeat_region   10859..11196
                /rpt_family="Alu"
repeat_region   11437..11614
                /rpt_family="L1"
repeat_region   11693..11935
                /rpt_family="L1"
repeat_region   12027..12252
                /rpt_family="L1"
repeat_region   12335..12445
                /rpt_family="MaLR"
repeat_region   12556..12640
                /rpt_family="L1"
repeat_region   13184..13471
                /rpt_family="B4"
repeat_region   14659..15026
                /rpt_family="L1"
repeat_region   15775..16125
                /rpt_family="MaLR"
repeat_region   16158..16313
                /rpt_family="L1"
repeat_region   16314..16541
                /rpt_family="MaLR"
repeat_region   16587..17116
                /rpt_family="L1"
repeat_region   17143..17256
                /rpt_family="L1"
repeat_region   17352..17578
                /rpt_family="L1"
repeat_region   18997..19181
                /rpt_family="L2"
repeat_region   19932..20138
                /rpt_family="B2"
repeat_region   20186..20395
                /rpt_family="ERV1"
repeat_region   20418..20567
                /rpt_family="B2"
repeat_region   21808..21855
                /rpt_family="Alu"
repeat_region   21964..22302
                /rpt_family="L1"
repeat_region   22999..23076
                /rpt_family="L1"
repeat_region   23077..23150
                /rpt_family="B4"
repeat_region   23211..24022
                /rpt_family="ID"
repeat_region   24086..24219
                /rpt_family="RMER1B"
repeat_region   25328..25749
                /rpt_family="Alu"
repeat_region   25961..26285
                /rpt_family="MaLR"
repeat_region   26445..26691
                /rpt_family="MaLR"
repeat_region   26695..26951
                /rpt_family="L1"
repeat_region   27000..27172
                /rpt_family="L1"
repeat_region   27010..27203
                /rpt_family="B4"
repeat_region   27578..27807
                /rpt_family="B2"

repeat_region   27865..28310
                /rpt_family="RMER15"
repeat_region   28497..28549
                /rpt_family="L1"
repeat_region   28554..29041
                /rpt_family="ERVL"
repeat_region   29627..29903
                /rpt_family="ERVK"
repeat_region   29924..30062
                /rpt_family="L1"
repeat_region   30385..30438
                /rpt_family="B4"
repeat_region   30522..30797
                /rpt_family="ID"
repeat_region   30977..31115
                /rpt_family="MaLR"
repeat_region   31160..31373
                /rpt_family="Alu"
repeat_region   31454..31544
                /rpt_family="L1"
repeat_region   32135..32285
                /rpt_family="L1"
repeat_region   32888..32959
                /rpt_family="B4"
repeat_region   33633..33736
                /rpt_family="Alu"
repeat_region   33974..34383
                /rpt_family="Alu"
repeat_region   35030..35109
                /rpt_family="ERVK"
repeat_region   35110..35394
                /rpt_family="RMER19B"
```

```
Query Match             13.6%;  Score 293.6;  DB 10;  Length 188791;
Best Local Similarity   83.4%;  Pred. No. 3.4e-58;
Matches   357;  Conservative   0;  Mismatches  69;  Indels  2;  Gaps  2;

Qy     429  GTTTCCTGAGGTTGTTCCCCTTAACATCGGAGGGGCTCACTTCACTACACGCCTGTCCAC  488
            ||||||| ||||||||||| |||||||| ||||||||||||||||| ||| |||||||
Db   96206  GTTTCCTGAGTTCGTCCTCCTTAACATTGGAGGGGCTCACTTTACCACGTGATAGTCTAC  96147

Qy     489  ACTGCGGTGCTACGAAGACACCATGTTGGCAGCGCCATGTTCAGTGGCGGCACTACATCCC  548
            |||| |||| |||||||| ||||| ||| ||| ||||||||| || |||||| ||||||||
Db   96146  TCTGCAGCACTATGAAGACA-CATGCTGGCTGTCATGTTTAGAGGGCGGCATTACATCCC  96088

Qy     549  CACAGACTCCGAGGGCCCGGTACTTCATCGACCGAGATGGCACACACTTTGGAGATGTGCT  608
            ||||||||| ||| ||| ||||||||||| || ||| |||| |||||||||||||| |||
Db   96087  TACAGACTCCAAGGATTGGTACTTCATCAACAGACGACCACACACTTTGGAGATGCGCT  96028

Qy     609  GAATTCCTGCGCTCAGGGAGCCTCCACCCAGGGAGCGTGTTCGAGCTGTGTACAAAGA  668
            |||  ||||||| |||||||||||| ||||||||||||| ||||||| ||||||||||
Db   96027  GAACTTCCTGCGCATTCAGGGAGCCTGCCACCCGGGGAGCAAGTGCGAGCTGTGCACAAAGA  95968

Qy     669  GGCCCAGTACTATGCCATGGGGCCCCCTCCTGAGCAGCGCCCTGAGAACATGCAGCCACTGAA  728
            || |||||||||||||| ||  ||||||||||||||||   |||||||||||||||   |||
Db   95967  GGTCCAGTACTATGCCATTGGACCCCCTCCTGAGCAGCGTGGAGAACATGCAGCCACCCGAA  95908

Qy     729  GGGCGGAGAAGTGCGCCAAGCGTTTCTGGGGACTCATGCCCATTACAAAGACCACTTGGA  788
            ||| | ||||||||||||| | ||||||||| ||||||  ||||||||| ||||| ||||
Db   95907  GGGTGAGAAGGTGCGCCAGACCTTCCTGTGGGCTCATGTCC-ATTACAAGGACCATTTGGA  95849

Qy     789  GCGGATTGTGGAGATCGGCCCCGGCTGCCGGTCCAGCGGAGGCCCGCTTTGCCAAGCT  848
            || ||| |||| | |   ||| | ||||  |||||| ||| ||| ||||||||||||||
Db   95848  GCCGGATAGTGGCGCGATTGCCCAGCTTCATCAGATACAGCGGAAGGCCCGCTTTGCCAAGCT  95789

Qy     849  CAAGAGCT  856
            || |||
Db   95788  GAAAGTCT  95781

RESULT 14
AC122830/c
LOCUS       AC122830          257003 bp     DNA        linear     ROD 08-NOV-2003
```

```
DEFINITION   Mus musculus BAC clone RP23-224I23 from 8, complete sequence.
ACCESSION    AC122830
VERSION      AC122830.4   GI:29469632
KEYWORDS     HTG.
SOURCE       Mus musculus (house mouse)
  ORGANISM   Mus musculus
             Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
             Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE    1  (bases 1 to 257003)
  AUTHORS    Scott,K., Harkins,R., Cotton,M., Spalding,L., Creason,K.,
             Mangiapanello,L. and Delaney,K.
  TITLE      The sequence of Mus musculus BAC clone RP23-224I23
  JOURNAL    Unpublished (2001)
REFERENCE    2  (bases 1 to 257003)
  AUTHORS    Wilson,R.
  TITLE      Sequencing of Mus musculus
  JOURNAL    Unpublished (2001)
REFERENCE    3  (bases 1 to 257003)
  AUTHORS    McPherson,J.D. and Waterston,R.H.
  TITLE      Direct Submission
  JOURNAL    Submitted (25-MAY-2002) Genome Sequencing Center, 4444 Forest Park
             Parkway, St. Louis, MO 63108, USA
REFERENCE    4  (bases 1 to 257003)
  AUTHORS    McPherson,J.D. and Waterston,R.H.
  TITLE      Direct Submission
  JOURNAL    Submitted (01-SEP-2002) Genome Sequencing Center, 4444 Forest Park
             Parkway, St. Louis, MO 63108, USA
REFERENCE    5  (bases 1 to 257003)
  AUTHORS    McPherson,J.D. and Waterston,R.H.
  TITLE      Direct Submission
  JOURNAL    Submitted (02-APR-2003) Genome Sequencing Center, 4444 Forest Park
             Parkway, St. Louis, MO 63108, USA
REFERENCE    6  (bases 1 to 257003)
  AUTHORS    Wilson,R.
  TITLE      Direct Submission
  JOURNAL    Submitted (08-NOV-2003) Department of Genetics, Washington
             University, 4444 Forest Park Avenue, St. Louis, Missouri 63108, USA
COMMENT      On Apr 2, 2003 this sequence version replaced gi:22539363.
             --------------  Genome Center
             Center: Washington University Genome Sequencing Center
             Center code: WUGSC
             Web site: http://genome.wustl.edu
             Contact: submissions@watson.wustl.edu
             --------------  Summary Statistics
             Center project name: M_BA0224I23
             --------------

             NOTICE: This sequence may not represent the entire insert of this
             clone.  It may be shorter because we only sequence overlapping
             clone sections once, or longer because we provide a small overlap
             between neighboring data submissions.

             This sequence was finished as follows unless otherwise noted:
             all regions were double stranded, sequenced with an alternate
             chemistry, or covered by high quality data (i.e., phred quality >=
             30); an attempt was made to resolve all sequencing problems, such
             as compressions and repeats; all regions were covered by sequence
             from more than one subclone; and the assembly was confirmed by
             restriction digest.

             MAPPING INFORMATION:
             Mapping information for this clone was provided by Dr. Wes Warren,
             Department of Genetics, Washington University, St. Louis MO. For
             additional information about the map position of this sequence, see
             http://genome.wustl.edu

             SOURCE INFORMATION:
             The RPCI-23 BAC Library has been constructed by Kazutoyo Osegawa
             and Minako Tateno in the laboratory of Pieter de Jong
             (http://www.chori.org) from female C57BL/6J mouse kidney and/or
             brain genomic DNA. The clone and detailed information can be
             obtained from Research Genetics, Inc. (http://www.resgen.com) or
             Pieter de Jong and coworkers at http://www.chori.org

NEIGHBORING SEQUENCE INFORMATION:
             This sequence is the entire insert of the clone. This clone is
             overlapped by AC117185 and AC116323.
FEATURES             Location/Qualifiers
     source          1..257003
                     /organism="Mus musculus"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:10090"
                     /chromosome="8"
                     /map="8"
                     /clone="RP23-224I23"
                     /clone_lib="RPCI-23"
     repeat_region   1..812
                     /rpt_family="L1"
     repeat_region   923..1089
                     /rpt_family="B2"
     tRNA            complement(1012..1087)
                     /product="tRNA-Ser"
                     /note="likely pseudogene (HMM Sc=18.64 / Sec struct
                     Sc=4.38)"
     repeat_region   1344..1615
                     /rpt_family="MaLR"
     repeat_region   1669..1779
                     /rpt_family="L1"
     repeat_region   1874..1967
                     /rpt_family="L1"
     repeat_region   3001..3109
                     /rpt_family="Alu"
     repeat_region   3187..3385
                     /rpt_family="B2"
     repeat_region   3396..3594
                     /rpt_family="B2"
     repeat_region   3602..3724
                     /rpt_family="BC1_MM"
     repeat_region   3727..3832
                     /rpt_family="Alu"
     repeat_region   4558..5185
                     /rpt_family="L1"
     repeat_region   5186..5331
                     /rpt_family="MaLR"
     repeat_region   5430..6192
                     /rpt_family="L1"
     repeat_region   6291..6790
                     /rpt_family="L1"
     repeat_region   6806..7044
                     /rpt_family="L1"
     repeat_region   7821..7933
                     /rpt_family="B2"
     repeat_region   8266..8352
                     /rpt_family="L2"
     repeat_region   8372..9045
                     /rpt_family="L1"
     repeat_region   9835..9956
                     /rpt_family="L2"
     repeat_region   9976..10076
                     /rpt_family="B4"
     repeat_region   10351..10966
                     /rpt_family="Alu"
     repeat_region   11410..11602
                     /rpt_family="MaLR"
     repeat_region   11622..11816
                     /rpt_family="B2"
     repeat_region   12435..12625
                     /rpt_family="L1"
     repeat_region   12771..12879
                     /rpt_family="L1"
     repeat_region   12882..13609
                     /rpt_family="Alu"
     repeat_region   13610..13823
                     /rpt_family="L1"
     repeat_region   13830..14073
                     /rpt_family="L1"
```

```
repeat_region    14361..14478
                 /rpt_family="Alu"
repeat_region    14542..14741
                 /rpt_family="B2"
repeat_region    14813..15193
                 /rpt_family="MaLR"
repeat_region    15194..15362
                 /rpt_family="MaLR"
repeat_region    15425..15923
                 /rpt_family="MaLR"
repeat_region    16265..16302
                 /rpt_family="MaLR"
repeat_region    16384..16540
                 /rpt_family="B4"
repeat_region    16597..16704
                 /rpt_family="Alu"
repeat_region    16785..16992
                 /rpt_family="B2"
repeat_region    20606..21525
                 /rpt_family="L1"
repeat_region    21523..21800
                 /rpt_family="L1"
repeat_region    21801..22146
                 /rpt_family="MaLR"
repeat_region    22147..22167
                 /rpt_family="L1"
repeat_region    22359..23677
                 /rpt_family="L1"
repeat_region    23820..23966
                 /rpt_family="L1"
repeat_region    23973..24696
                 /rpt_family="Alu"
repeat_region    24695..25435
                 /rpt_family="L1"
repeat_region    25441..25940
                 /rpt_family="L1"
repeat_region    26315..27000
                 /rpt_family="L1"
repeat_region    27001..27373
                 /rpt_family="L1"
repeat_region    27374..27429
                 /rpt_family="MaLR"
repeat_region    27640..27841
                 /rpt_family="L1"
repeat_region    27880..28183
                 /rpt_family="ERVK"
repeat_region    28861..28977
                 /rpt_family="L1"
repeat_region    29050..29397
                 /rpt_family="MaLR"
repeat_region    29504..29537
                 /rpt_family="MaLR"
repeat_region    29550..29685
                 /rpt_family="Alu"
```

Query Match   13.6%;   Score 293.6; DB 10; Length 257003;
Best Local Similarity   83.4%;   Pred. No. 3.6e-58;
Matches 357;   Conservative   0;   Mismatches 69;   Indels   2;   Gaps   2;

```
Qy    429 GTTTCCTGAGGTTGTTCCCTTAACATCGGAGGGGCTCACTTCACTACACGCCTGTCCAC 488
          ||||| |||||||||| |||||||||||||||| ||||||||| ||||||||||||||
Db  55806 GTTTCCTGAGTTCGTCTCCTTAACATTGGAGGGGCTCACTTTACCACGTGATAGTCTAC 55747

Qy    489 ACTGGCGGTGCTACGAAGACACCATGTTGTGGCAGCCATGTTCAGTGGGGGCACTACATCCC 548
          ||||| |||||| ||||||| |||||||||||||||||||||||||| ||||||||||||||
Db  55746 TCTGCAGCACTATGAAGACA-CATGCTGGCTGTCGCTGCTCATGTTTAGAGGGCCATTACATCCC 55688

Qy    549 CACAGACTCCGAGGCCGTACTTCATCGACCGAGATGGCACACACTTTGGAGATGTGCT 608
          ||||||||||| |||||| ||||||| ||||||||||||||||||||||||||||| ||
Db  55687 TACAGACTCCAAGGATTGGTACTTCATCAACAGACAGCACACACTTTGGAGATGGCGCT 55628

Qy    609 GAATTTCCTGCGCTCAGGGGACCTCCCACCCCCAGGGAGCCGTGTTCGAGCTGTGTACAAAGA 668
          ||||| |||| |||||||| |||||||||||||||||||||||||||||||||||||| |||||
Db  55627 GAACTTCCTGCATTCAGGGGACCCTGCCACCCTGCCACCCGGGGAGCAAGTGCGAGCTGTGCACAAAGA 55568
```

```
Qy    669 GGGCCAGTACTATGCCATCGGCGCCCCTCCTGGAGCAGCTGAGAGAACATGCAGCCACTGAA 728
          |||||||||||||||||||| |||||||||||||||||||||||||||||||||| |||||||
Db  55567 GGTTCCAGTACTATGCCATTGCCACCCCTCCTGGAGCAGCTGGAGAACATGCAGAACCAGCCACCGAA 55508

Qy    729 GGGCGAGAAGTGCGCCAAGGTTCTGGAGACTCATGCGCCTTATTACACAGAGACCACTTGGA 788
          ||||||||||||||| ||||||||||||| ||||| ||| |||||||||| |||||||||||
Db  55507 GGGTGAGAAGTGCGCCAGACCTTCCTGGGGCTCATGTCC-ATTACAAGAGACCATTTGGA 55449

Qy    789 GCGGATTGTGGAGATCGCCCCGGCGTGCCGTCCAGCGGGAAGCCCGCTTTGCCAAGCT 848
          ||||| |||||||| |||||||||||||| ||||||| ||||||||||||||||||||
Db  55448 GCGGATAGTGGCGATTGCCCAGCTTCATACAGCGAAGCGGAAGGCCCGCTTTGCCAAGCT 55389

Qy    849 CAAGAGCT 856
          ||||  ||
Db  55388 GAAAGTCT 55381
```

RESULT 15
BC051544

```
LOCUS       BC051544                1103 bp    mRNA    linear    ROD 04-NOV-2003
DEFINITION  Mus musculus potassium channel tetramerisation domain containing 7,
            mRNA (cDNA clone IMAGE:5254261), partial cds.
ACCESSION   BC051544
VERSION     BC051544.1    GI:30186133
KEYWORDS
SOURCE      Mus musculus (house mouse)
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 1103)
  AUTHORS   Strausberg,R.L., Feingold,E.A., Grouse,L.H., Derge,J.G.,
            Klausner,R.D., Collins,F.S., Wagner,L., Shenmen,C.M., Schuler,G.D.,
            Altschul,S.F., Zeeberg,B., Buetow,K.H., Schaefer,C.F., Bhat,N.K.,
            Hopkins,R.F., Jordan,H., Moore,T., Max,S.I., Wang,J., Hsieh,F.,
            Diatchenko,L., Marusina,K., Farmer,A.A., Rubin,G.M., Hong,L.,
            Stapleton,M., Soares,M.B., Bonaldo,M.F., Casavant,T.L.,
            Scheetz,T.E., Brownstein,M.J., Usdin,T.B., Toshiyuki,S.,
            Carninci,P., Prange,C., Raha,S.S., Loquellano,N.A., Peters,G.J.,
            Abramson,R.D., Mullahy,S.J., Bosak,S.A., McEwan,P.J.,
            McKernan,K.J., Malek,J.A., Gunaratne,P.H., Richards,S.,
            Worley,K.C., Hale,S., Garcia,A.M., Gay,L.J., Hulyk,S.W.,
            Villalon,D.K., Muzny,D.M., Sodergren,E.J., Lu,X., Gibbs,R.A.,
            Fahey,J., Helton,E., Ketteman,M., Madan,A., Rodrigues,S.,
            Sanchez,A., Whiting,M., Madan,A., Young,A.C., Shevchenko,Y.,
            Bouffard,G.G., Blakesley,R.W., Touchman,J.W., Green,E.D.,
            Dickson,M.C., Rodriguez,A.C., Grimwood,J., Schmutz,J., Myers,R.M.,
            Butterfield,Y.S., Krzywinski,M.I., Skalska,U., Smailus,D.E.,
            Schnerch,A., Schein,J.E., Jones,S.J. and Marra,M.A.
  TITLE     Generation and initial analysis of more than 15,000 full-length
            human and mouse cDNA sequences
  JOURNAL   Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)
  MEDLINE   22388257
  PUBMED    12477932
REFERENCE   2  (bases 1 to 1103)
  AUTHORS   Strausberg,R.
  TITLE     Direct Submission
  JOURNAL   Submitted (28-APR-2003) National Institutes of Health, Mammalian
            Gene Collection (MGC), Cancer Genomics Office, National Cancer
            Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
            USA
  REMARK    NIH-MGC Project URL: http://mgc.nci.nih.gov
  COMMENT   Contact: MGC help desk
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: Gilbert Smith, Ph.D.
            cDNA Library Preparation: M. Bento Soares, University of Iowa
            cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
            DNA Sequencing by: Institute for Systems Biology
            http://www.systemsbiology.org
            contact: amadan@systemsbiology.org
            Anup Madan, Jessica Fahey, Erin Helton, Mark Ketteman, Anuradha
            Madan, Stephanie Rodrigues, Amy Sanchez and Michelle Whiting
```

```
          Clone distribution: MGC clone distribution information can be found
          through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov
          Series: IRAK Plate: 113 Row: 1 Column: 15.
FEATURES            Location/Qualifiers
     source         1..1103
                    /organism="Mus musculus"
                    /mol_type="mRNA"
                    /strain="Czech II"
                    /db_xref="taxon:10090"
                    /clone="IMAGE:5254261"
                    /tissue_type="Lung, 2 pooled mouse tumors."
                    /clone_lib="NCI_CGAP_Lu33"
                    /lab_host="DH10B"
                    /note="Vector: pT7T3-Pac"
     gene           <1..1103
                    /gene="Kctd7"
                    /note="synonym: 4932409E18"
                    /db_xref="LocusID:212919"
                    /db_xref="MGI:2442265"
     CDS            <1..608
                    /gene="Kctd7"
                    /codon_start=3
                    /product="Kctd7 protein"
                    /protein_id="AAH51544.1"
                    /db_xref="GI:30186134"
                    /db_xref="LocusID:212919"
                    /db_xref="MGI:2442265"
                    /translation="HEAEGRYFIDRDGTHFGDVLNFLRSGDLPPREHVRAVHKEAQYY
                    AIGPLLEQLENMQPLKGEKVRQAFLGLMPYYKDHLERIVEIARLRAVQRKARFAKLKV
                    CVFKEEMPITPYECPLLNSLRFERSESDGQLFEHHCEVDVSFGPWEAVADVYDLLHCL
                    VTDLSAQGLITVDHQCIGVCDKHLVNHYCKRPIYEFKITWW"
     misc_feature   6..152
                    /gene="Kctd7"
                    /note="K_tetra; Region: K+ channel tetramerisation domain.
                    The N-terminal, cytoplasmic tetramerisation domain (T1) of
                    voltage-gated K+ channels encodes molecular determinants
                    for subfamily-specific assembly of alpha-subunits into
                    functional tetrameric channels. It is distantly related to
                    the BTB/POZ domain pfam00651"
                    /db_xref="CDD:pfam02214"

ORIGIN

Query Match         12.5%;    Score 268.8;  DB 10;  Length 1103;
Best Local Similarity  92.8%;    Pred. No. 1.2e-52;
Matches 282;   Conservative   0;  Mismatches  22;   Indels   0;   Gaps    0;

Qy    553 GACTCCGAGGGCCGGTACTTCATCGACCGAGATGGCACACACTTGGAGATGTGCTGAAT 612
              ||||||||| ||||||||||||| | ||||||||||  || ||||||||| |||||||
Db      6 GAGGCCGAGGGCCCGTACTTCATCGATCGAGATCGACACACACTTTGGAGATGTGCTGAAC 65

Qy    613 TTCCTGCGCTCAGGGGACCTCCCCACCCAGGGAGCTGTTCGAGCTGTGTACAAAGAGGCC 672
              |||||||||||||||| |||  | |||| || |  || |||||||||||||||||||||
Db     66 TTCCTGCGTTCAGGGGACCTGCCCCCCAGGGAGCATGTGCGAGCTGTGCACAAGGAGGCC 125

Qy    673 CAGTACTATGCCATCGGGCCCCTCTGGAGCAGCTGAGAACATGCAGCCACTGAAGGGC 732
              |||||||||||||||||| ||||| ||||||||| ||||||||||||||||||||| |||
Db    126 CAGTACTATGCCATCGGGCCCCTCTGGAGCAGCTGAGAACATGCAGCCACTGAAGGGC 185

Qy    733 GAGAAGGTGCGCCAAGCGTTTCTGGGACTCATCGCCTATTACAAAGACCACTTGGAGCGG 792
              |||||||||| |||| || ||||| || ||||| |||||||||||| |||||||||||
Db    186 GAGAAGGTGCGCCAGGCCCTTCCTGGGACTCATGCCCTATTACAAAGACCACTTGGAGCGG 245

Qy    793 ATTGTGGAGATCGCCCGGCTGCGCCGGCTGCTGGTCCAGGGAGCTCCGTTTGCCAAGCTCAAG 852
              |||||||||||| || ||||| | |||| ||| |||||||||||||||||||||||||||
Db    246 ATTGTGGAGATTGCCCGGCTTCGCCGCAGCACCGGAGCCGGAGGCCCGCTTTGCCAAGCTCAAA 305

Qy    853 AGCT 856
              || |
Db    306 GTCT 309

Search completed: October  8, 2004, 08:34:53
Job time : 8355 secs
```

GenCore version 5.1.6

Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: October 8, 2004, 08:34:58 ; Search time 1029 Seconds
(without alignments)
10612.096 Million cell updates/sec

Title: US-10-086-156-23
Perfect score: 2154
Sequence: 1 atgacgatggcggtttgcg........taataaaaaaaaaaaa 2154

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 3340653 seqs, 2534783454 residues

Total number of hits satisfying chosen parameters: 6681306

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published Applications NA:*
1: /cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq:*
2: /cgn2_6/ptodata/1/pubpna/PCT_NEW_PUB.seq:*
3: /cgn2_6/ptodata/1/pubpna/US06_NEW_PUB.seq:*
4: /cgn2_6/ptodata/1/pubpna/US06_PUBCOMB.seq:*
5: /cgn2_6/ptodata/1/pubpna/US07_NEW_PUB.seq:*
6: /cgn2_6/ptodata/1/pubpna/PCTUS_PUBCOMB.seq:*
7: /cgn2_6/ptodata/1/pubpna/US08_NEW_PUB.seq:*
8: /cgn2_6/ptodata/1/pubpna/US08_PUBCOMB.seq:*
9: /cgn2_6/ptodata/1/pubpna/US09A_PUBCOMB.seq:*
10: /cgn2_6/ptodata/1/pubpna/US09B_PUBCOMB.seq:*
11: /cgn2_6/ptodata/1/pubpna/US09C_PUBCOMB.seq:*
12: /cgn2_6/ptodata/1/pubpna/US09_NEW_PUB.seq:*
13: /cgn2_6/ptodata/1/pubpna/US09_NEW_PUB.seq:*
14: /cgn2_6/ptodata/1/pubpna/US10A_PUBCOMB.seq:*
15: /cgn2_6/ptodata/1/pubpna/US10B_PUBCOMB.seq:*
16: /cgn2_6/ptodata/1/pubpna/US10C_PUBCOMB.seq:*
17: /cgn2_6/ptodata/1/pubpna/US10_NEW_PUB.seq:*
18: /cgn2_6/ptodata/1/pubpna/US10_NEW_PUB.seq:*
19: /cgn2_6/ptodata/1/pubpna/US60_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.		Score	% Query Match	Length	DB	ID		Description
1		2154	100.0	2154	15	US-10-086-156-23		Sequence 23, Appl
2		587	27.3	2576	16	US-10-094-749-1045		Sequence 1045, Ap
3		585.4	27.2	1068	16	US-10-120-988-408		Sequence 408, App
4		465	21.6	519	13	US-10-296-115-18		Sequence 18, Appl
5		417	19.4	1100	16	US-10-027-632-31358		Sequence 31358, A
6		417	19.4	1100	16	US-10-027-632-31358		Sequence 31358, A
7		387.8	18.0	680	14	US-10-024-579-8		Sequence 8, Appli
8		315.2	14.6	363	14	US-10-024-579-4		Sequence 4, Appli
9	c	174	8.1	473	9	US-09-864-761-1385		Sequence 1385, Ap
10	c	174	8.1	583	15	US-10-086-156-29		Sequence 29, Appl
11		172.4	8.0	321	14	US-10-024-579-6		Sequence 6, Appli
12	c	170.4	7.9	173	9	US-09-864-761-18144		Sequence 18144, A
13		116.4	5.4	18534	13	US-10-240-425-1575		Sequence 1575, Ap
14	c	116.4	5.4	60430	13	US-10-087-192-1732		Sequence 1732, Ap
15	c	116.2	5.4	646	13	US-10-027-632-110966		Sequence 110966,
16	c	116.2	5.4	646	13	US-10-027-632-110967		Sequence 110967,
17	c	116.2	5.4	646	13	US-10-027-632-110968		Sequence 110968,
18	c	116.2	5.4	646	16	US-10-027-632-110966		Sequence 110966,
19	c	116.2	5.4	646	16	US-10-027-632-110967		Sequence 110967,
20	c	116.2	5.4	646	16	US-10-027-632-110968		Sequence 110968,
21		111	5.2	535	16	US-10-027-632-281283		Sequence 281283,
22		111	5.2	535	16	US-10-027-632-281283		Sequence 281283,
23	c	111	5.2	43871	17	US-10-741-601-5628		Sequence 5628, Ap
24	c	110.8	5.1	804	15	US-10-198-846-4375		Sequence 4375, Ap
25		110.6	5.1	535	13	US-10-027-632-281282		Sequence 281282,
26		110.6	5.1	535	16	US-10-027-632-281282		Sequence 281282, A
27		110.4	5.1	2094	13	US-10-027-632-99288		Sequence 99288, A
28		110.4	5.1	2094	13	US-10-027-632-99289		Sequence 99289, A
29		110.4	5.1	2094	16	US-10-027-632-99288		Sequence 99288, A
30		110.4	5.1	2094	16	US-10-027-632-99289		Sequence 99289, A
31		110	5.1	87869	17	US-10-717-597-194		Sequence 194, App
32		109.8	5.1	1595	9	US-09-822-849A-250		Sequence 250, App
33		109.8	5.1	2027	16	US-10-104-047-100		Sequence 100, App
34		109.4	5.1	227931	16	US-10-085-117-274		Sequence 274, App
35	c	109.4	5.1	325348	16	US-10-085-117-358		Sequence 358, App
36	c	109.2	5.1	28667	16	US-10-264-237-2863		Sequence 2863, Ap
37	c	109.2	5.1	28668	16	US-10-264-237-2864		Sequence 2864, Ap
38	c	109	5.1	80420	17	US-10-322-281-42		Sequence 42, Appl
39		106.8	5.0	168198	17	US-10-322-696-88		Sequence 88, Appl
40		106.2	4.9	1246	13	US-10-027-632-122821		Sequence 122821,
41		106.2	4.9	1246	13	US-10-027-632-122822		Sequence 122822,
42		106.2	4.9	1246	16	US-10-027-632-122821		Sequence 122821,
43		106.2	4.9	1246	16	US-10-027-632-122822		Sequence 122822,
44		106.2	4.9	3139	13	US-10-027-632-115236		Sequence 115236,
45		106.2	4.9	3139	13	US-10-027-632-115237		Sequence 115237,

ALIGNMENTS

RESULT 1
US-10-086-156-23
; Sequence 23, Application US/10086156
; Publication No. US20030054989A1
; GENERAL INFORMATION:
; APPLICANT: Bristol-Myers Squibb Company
; TITLE OF INVENTION: POLYNUCLEOTIDE ENCODING TWO NOVEL HUMAN POTASSIUM CHANNEL BETA-SU
; TITLE OF INVENTION: K+betaM4 and K+betaM5
; FILE REFERENCE: D015NP
; CURRENT APPLICATION NUMBER: US/10/086,156
; CURRENT FILING DATE: 2002-02-28
; PRIOR APPLICATION NUMBER: US 60/272,190
; PRIOR FILING DATE: 2001-02-28
; PRIOR APPLICATION NUMBER: US 60/274,258
; PRIOR FILING DATE: 2001-03-07
; NUMBER OF SEQ ID NOS: 98
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 23
; LENGTH: 2154
; TYPE: DNA
; ORGANISM: homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(1029)
US-10-086-156-23

Query Match   100.0%; Score 2154; DB 15; Length 2154;
Best Local Similarity   100.0%; Pred. No. 0;
Matches 2154; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy   1 ATGACGATGGCGGTTTGCGGAATAGAAAAGGGGGAAAGGGACCACCTCAGGGCGCGGCCG 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1 ATGACGATGGCGGTTTGCGGAATAGAAAAGGGGGAAAGGGACCACCTCAGGGCGCGGCCG 60

Qy   61 CTGGCGCTGCCTGCTCTTCGACTGGCGAGCTTCCTGCCAATCAGGGGCGGAACCAGCGCG 120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   61 CTGGCGCTGCCTGCTCTTCGACTGGCGAGCTTCCTGCCAATCAGGGGCGGAACCAGCGCG 120

```
Qy   121  GCGTCGGCCAGTAGCGGGAGGCGGTCGGGTCGGGCGGGAGCGGTCGGC  180
          ||||||||||||||||||||||||||||||||||||||||||||||||
Db   121  GCGTCGGCCAGTAGCGGGAGGCGGTCGGGTCGGGCGGGAGCGGTCGGC  180

Qy   181  GTTGAGGGAGCCACCCGCCCCTCCCGCCACTGCGCTCTCCGCCGCCC   240
          |||||||||||||||||||||||||||||||||||||||||||||||
Db   181  GTTGAGGGAGCCACCCGCCCCTCCCGCCACTGCGCTCTCCGCCGCCC   240

Qy   241  GCAGCCGGCCCGCGGCGCTCATGCGGATGCCAGGAGCTCGGCGAGCCAGGAGGATGCGTGGTAGTCACG  300
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   241  GCAGCCGGCCCGCGGCGCTCATGCGGATGCCAGGAGCTCGGCGAGCCAGGAGGATGCGTGGTAGTCACG  300

Qy   301  GGGCGGGAGCCAGACAGCCGTCGTCAGGACGGTCGTGCCATGTCCAGCTCGACGCCGAAGAC  360
          |||||||||||||||||||||||||||||||||||||||||||||||||||
Db   301  GGGCGGGAGCCAGACAGCCGTCGTCAGGACGGTCGTGCCATGTCCAGCTCGACGCCGAAGAC  360

Qy   361  GACTTTCTGGAGCCGGCCGGCCACGCGGCCACGCCGCAGCGGGCGGGCCTGCCCTGCTG  420
          |||||||||||||||||||||||||||||||||||||||||||||
Db   361  GACTTTCTGGAGCCGGCCGGCCACGCGGCCACGCCGCAGCGGGCGGGCCTGCCCTGCTG  420

Qy   421  CCACAGAGTTTCCTGAGGTTGTTCCCCTTAACATCGGAGGGCTCACTTCACTACACGC  480
          |||||||||||||||||||||||||||||||||||||||||||||||
Db   421  CCACAGAGTTTCCTGAGGTTGTTCCCCTTAACATCGGAGGGCTCACTTCACTACACGC  480

Qy   481  CTGTCCACACTGCGGTGCTACGAAGACACCATGTTGGCAGCCATGTTCAGTGGGCGGCGCAC  540
          ||||||||||||||||||||||||||||||||||||||||||||||||||
Db   481  CTGTCCACACTGCGGTGCTACGAAGACACCATGTTGGCAGCCATGTTCAGTGGGCGGCGCAC  540

Qy   541  TACATCCCCACAGACTCCGAGGGCCGGCCGGTACTTCATCGACCGAGATGCGACACCACTTGGA  600
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   541  TACATCCCCACAGACTCCGAGGGCCGGCCGGTACTTCATCGACCGAGATGCGACACCACTTGGA  600

Qy   601  GATGTGCTGAATTCCTGCGCTCAGGGGAGCCTGTTCGAGCTGTG  660
          ||||||||||||||||||||||||||||||||||||||||||
Db   601  GATGTGCTGAATTCCTGCGCTCAGGGGAGCCTGTTCGAGCTGTG  660

Qy   661  TACAAAGAGGCCCAGTACTATCGGCCATCGGGCCCCCTCCTGGAGCAGTGGAGAACATGCAG  720
          |||||||||||||||||||||||||||||||||||||||||||||||||||
Db   661  TACAAAGAGGCCCAGTACTATCGGCCATCGGGCCCCCTCCTGGAGCAGTGGAGAACATGCAG  720

Qy   721  CCACTGAAGAGGCGAGAAGGTGCGCCAAGCGTTCGGGACTCATGCCAAGAC  780
          |||||||||||||||||||||||||||||||||||||||||
Db   721  CCACTGAAGAGGCGAGAAGGTGCGCCAAGCGTTCGGGACTCATGCCAAGAC  780

Qy   781  CACTTGGAGCGGGATTGTGAGAGTCGCGGTCGCCAGCCGGAAGGCCCGCTTT  840
          ||||||||||||||||||||||||||||||||||||||||||
Db   781  CACTTGGAGCGGGATTGTGAGAGTCGCGGTCGCCAGCCGGAAGGCCCGCTTT  840

Qy   841  GCCAAGCTCAAGAGCTTGACACCTTCCTGGCTAATGAGTGTCCTCATCAAGATGCCCCCT  900
          |||||||||||||||||||||||||||||||||||||||||||||||||
Db   841  GCCAAGCTCAAGAGCTTGACACCTTCCTGGCTAATGAGTGTCCTCATCAAGATGCCCCCT  900

Qy   901  GGAGTCACACATCATGGATTAACGGAAGAGGCCGCTGTATTTGGAAACTCCCATTGTCCA  960
          ||||||||||||||||||||||||||||||||||||||||||||||||
Db   901  GGAGTCACACATCATGGATTAACGGAAGAGGCCGCTGTATTTGGAAACTCCCATTGTCCA  960

Qy   961  GAGAGACAGAACAATGAGAAGAAATCCCCGTGTCCAGTTCAGGAGTATTCCAACAC  1020
          |||||||||||||||||||||||||||||||||||||||||||||
Db   961  GAGAGACAGAACAATGAGAAGAAATCCCCGTGTCCAGTTCAGGAGTATTCCAACAC  1020

Qy  1021  TTCATGGGCTAGAGGATTCCATTGAAGATGGGGTTTAGTCTTGATTTTGAACACCTGTCA  1080
          |||||||||||||||||||||||||||||||||||||||||||||||
Db  1021  TTCATGGGCTAGAGGATTCCATTGCATGGCAATCTGAACCCTTTTATGCCAACAACAC  1080

Qy  1081  GCACTGTTCTCTGTTGTTCATGCAATTCTGACCCTTTTATGCCAACAACACCCCTGGGAC  1140
          ||||||||||||||||||||||||||||||||||||||||||||||||
Db  1081  GCACTGTTCTCTGTTGTTCATGCAATTCTGACCCTTTTATGGCAACAACACCCCTGGGAC  1140

Qy  1141  AACCCAGATTTGTAGATTGAGATCCAAAGGTAGAATTTCCAGACAGTCCAACCAAGGTAT  1200
          ||||||||||||||||||||||||||||||||||||||||||||||||
Db  1141  AACCCAGATTTGTAGATTGAGATCCAAAGGTAGAATTTCCAGACAGTCCAACCAAGGTAT  1200
```

```
Qy  1201  CAAGTGATGTTCCAGAGTGGAAGGCTCTCACCGTGTCCCAGATTTCTGGGGTTTGTAA  1260
          ||||||||||||||||||||||||||||||||||||||||||||
Db  1201  CAAGTGATGTTCCAGAGTGGAAGGCTCTCACCGTGTCCCAGATTTCTGGGGTTTGTAA  1260

Qy  1261  GCAGTACTGGCCATTGTGACCCTGTTTTTTACCTATCATTCTGTCTTTTAGGACATG  1320
          |||||||||||||||||||||||||||||||||||||||||||||||
Db  1261  GCAGTACTGGCCATTGTGACCCTGTTTTTTACCTAATCATTCTGTCTTTTAGGACATG  1320

Qy  1321  GTTTTACCCGATCCTGCAAAGGATCCAGAATTCCAATAGCTGAAAACCCTGTTATAGC  1380
          |||||||||||||||||||||||||||||||||||||||||||||||
Db  1321  GTTTTACCCGATCCCTGGCAAAGGATCCAGAATTCCAATAGCTGAAAACCCTGTTATAGC  1380

Qy  1381  TTTTCTCCTATTCTGCCTTACCCAAGACACACTTGAACCCCTCAGTAAGGCTATAGAGAG  1440
          |||||||||||||||||||||||||||||||||||||||||||||||
Db  1381  TTTTCTCCTATTCTGCCTTACCCAAGACACACTTGAACCCCTCAGTAAGGCTATAGAGAG  1440

Qy  1441  GGCCATGAGCAGGGGCAGCCTCTCCCTGTTTCTACAGCTCCATGATGAGGGTTGACTG  1500
          |||||||||||||||||||||||||||||||||||||||||||||||
Db  1441  GGCCATGAGCAGGGGCAGCCCTCTCCCTGTTTCTACAGCTCCATGATGAGGGGTTGACTG  1500

Qy  1501  AGGCCAGCAATCCTTGTGAGGTCGACAGTTGCAATATAATTAACAGTTCAAGATCTAGA  1560
          |||||||||||||||||||||||||||||||||||||||||||||||
Db  1501  AGGCCAGCAATCCTTGTGAGGTCGACAGTTGCAATATAATTAACAGTTCAAGATCTAGA  1560

Qy  1561  GGTACCTTTGAAAGAACCCCTTCAGGGATATCATCCACAGTAGCCTGGAGCAGCCAAG  1620
          |||||||||||||||||||||||||||||||||||||||||||||||
Db  1561  GGTACCTTTGAAAGAACCCCTTCAGGGATATCATCCACAGTAGCCTGGAGCAGCCAAG  1620

Qy  1621  GTGAACCTGAGATTTGACCCACAATAAGGGGGCCATTCTTTTTCAAATATTTTGG  1680
          ||||||||||||||||||||||||||||||||||||||||||||
Db  1621  GTGAACCTGAGATTTGACCCACAATAAGGGGGCCATTCTTTTTCAAATATTTTTGG  1680

Qy  1681  CTTCAGAATACACTTCATTACACATGCAAATATTGAGAGATTAACAGAAATTCCAGCTCT  1740
          |||||||||||||||||||||||||||||||||||||||||||||||
Db  1681  CTTCAGAATACACTTCATTACACATGCAAATATTGAGAGATTAACAGAAATTCCAGCTCT  1740

Qy  1741  TATGCCTAACTGAGAAGAGCCACTGCAAGTTGCAGTTGGTACCATGTTAGTGCAGCAGGC  1800
          |||||||||||||||||||||||||||||||||||||||||||||||
Db  1741  TATGCCTAACTGAGAAGAGCCACTGCAAGTTGCAGTTGGTACCATGTTAGTGCAGCAGGC  1800

Qy  1801  CAGCTGAATCCCAGAGCTTCCCAAAGTGGACACCAGCGGGACTATCCTGATGTCCAC  1860
          |||||||||||||||||||||||||||||||||||||||||||||
Db  1801  CAGCTGAATCCCAGAGCTTCCCAAAGTGGACACCAGCGGGACTATCCTGATGTCCAC  1860

Qy  1861  CCAAGAGGAAGATGAGCTGAGGCGCTCTTGCTCTGCCCAAATGCATCCCATGTCATT  1920
          |||||||||||||||||||||||||||||||||||||||||||||
Db  1861  CCAAGAGGAAGATGAGCTGAGGCGCTCTTGCTCTGCCCAAATGCATCCCATGTCATT  1920

Qy  1921  CACGTGTCACCCATTCAAAATAACATGGCATTCTTGTAATCCTGACATGTAAGA  1980
          ||||||||||||||||||||||||||||||||||||||||||
Db  1921  CACGTGTCACCCATTCAAAATAACATGGCATTCTTGGAACCTTGTAATCCTGACATGTAAGA  1980

Qy  1981  CCAGCCTACACATTGGGGTGGGTGCAGGGCTCACACTTTGAAG  2040
          |||||||||||||||||||||||||||||||||||||
Db  1981  CCAGCCTACACATTGGGGTGGGTGCAGGGCTCACACTTTGAAG  2040

Qy  2041  GCTGAGGGTGGGCAGATTGCTTGAGCACAGGAGTTCCAGACCAACATGGCCA  2100
          ||||||||||||||||||||||||||||||||||||||||||
Db  2041  GCTGAGGGTGGGCAGATTGCTTGAGCACAGGAGTTCCAGACCAACATGGCCA  2100

Qy  2101  AATCCTGTCTCTTCAAGAAATAAAATAATAATAAAAAAAAAAAA  2154
          ||||||||||||||||||||||||||||||||||||
Db  2101  AATCCTGTCTCTTCAAGAAATAAAATAATAATAAAAAAAAAAAA  2154
```

; APPLICANT: WAKAMATSU, AI
; APPLICANT: SATO, HIROYUKI
; APPLICANT: ISHII, SHIZUKO
; APPLICANT: YAMAMOTO, JUN-ICHI
; APPLICANT: ISONO, YUUKO
; APPLICANT: HIO, YURI
; APPLICANT: OTSUKA, KAORU
; APPLICANT: NAGAI, KEIICHI
; APPLICANT: IRIE, RYOTARO
; APPLICANT: TAMECHIKA, ICHIRO
; APPLICANT: SEKI, NAOHIKO
; APPLICANT: YOSHIKAWA, TSUTOMU
; APPLICANT: OTSUKA, MOTOYUKI
; APPLICANT: NAGAHARI, KENJI
; APPLICANT: MASUHO, YASUHIKO

; TITLE OF INVENTION: NOVEL FULL-LENGTH CDNA
; FILE REFERENCE: 084335/0160
; CURRENT APPLICATION NUMBER: US/10/094,749
; CURRENT FILING DATE: 2002-03-12
; PRIOR APPLICATION NUMBER: 60/350,435
; PRIOR FILING DATE: 2002-01-24
; PRIOR APPLICATION NUMBER: JP 2001-328381
; PRIOR FILING DATE: 2001-09-14
; NUMBER OF SEQ ID NOS: 3381
; SOFTWARE: PatentIn Ver. 2.1
;
; SEQ ID NO 1045
; LENGTH: 2576
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-094-749-1045

Query Match   27.3%; Score 587; DB 16; Length 2576;
Best Local Similarity   96.0%; Pred. No. 9.9e-167;
Matches   602; Conservative   0; Mismatches 25; Indels   0; Gaps   0;

Qy   230 CCGGCCAGCCGCGAGCCGCGCCGCGCTCATGCCGAGGCCAGGAGGGATGG 289
             |||| ||||||| || | | || |||||||||||||
Db   52 CCGGGGCCCGCCGCTCCGGCGAAGCCGCGCCCGCCACTGCCCCAGAGCCAGAGCCAGAGGGATGG 111

Qy   290 TGGTAGTCACGGGGCGGGAGCCCAGACAGCCGTCGTCAGGACGTGCCATGTCCAGCTCTG 349
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   112 TGGTAGTCACGGGGCGGGAGCCCAGACAGCCGTCGTCAGGACGTGCCATGTCCAGCTCTG 171

Qy   350 ACGCCGAAGACGACTTTCTGGAGCCGGCCACGCGGGGGCACGCGGC 409
          |||||||||||||||||||||||||||||||||||||||||||||
Db   172 ACGCCGAAGACGACTTTCTGGAGCCGGCCACGCGGCGGCGGGGCACGCGGC 231

Qy   410 TGCCCCTGCTGCCACAGGAGTTTCCTGAGGTTGTCCCCTTAACATCGGAGGGGCTCACT 469
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   232 TGCCCCTGCTGCCACAGGAGTTTCCTGAGGTTGTCCCCTTAACATCGGAGGGGCTCACT 291

Qy   470 TCACTACACGCCTGTCCACACTGCGGTGCTACGAAGACACCATGTTGGCAGCCATGTTCA 529
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   292 TCACTACACGCCTGTCCACACTGCGGTGCTACGAAGACACCATGTTGGCAGCCATGTTCA 351

Qy   530 GTGGGCGGCACTACATCCCCACAGACTCCGAGGGCCCGGTACTTCATCGACCGAGATGGCA 589
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   352 GTGGGCGGCACTACATCCCCACAGACTCCGAGGGCCGCTCAGGGACCTTCCCACCCAGGGAGCGTG 411

Qy   590 CACACTTTGGAGATGTGTACAAAGAGGGCGGAAGTGCGCCAAGCGTTTCTGGGACTTCATGCCCT 649
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   412 CACACTTTGGAGATGTGTACAAAGAGGGCCCAGTACTTCATCGACCGAGATGGCG 471

Qy   650 TTCGAGCTGTGTACAAAGAGGGCGGAAGTGCGCCAAGCGTTTCTGGGACTTCATGCCCT 709
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   472 TTCGAGCTGTGTACAAAGAGGGCCCAGTACTTCATCGACCGAGATGGCG 531

Qy   710 AGAACATGCAGCCACTGGAGCCGGATTGTGGAGATCGCCGTCGGCGTTCGCCGTCCAGCGGGA 769
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   532 AGAACATGCAGCCACTGGAGCCGGATTGTGGAGATCGCCGTCGGCGTTCGCCGTCCAGCGGGA 591

Qy   770 ATTACAAAGACCACTTGGAGCGCGGATTGTGGAGATCGCCCGGCGTGCCGGTCCAGCGGGA 829
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   592 ATTACAAAGACCACTTGGAGCGCGGATTGTGGAGATCGCCCGGCGTGCCGGTCCAGCGGGA 651

Qy   830 AGGCCCGCTTGCCAAGCTCAAGAGCT 856
          |||||||||||||||||||||||||
Db   652 AGGCCCGCTTGCCAAGCTCAAGGTCT 678

RESULT 3
US-10-120-988-408
; Sequence 408, Application US/10120988
; Publication No. US2003021974SA1
; GENERAL INFORMATION:
; APPLICANT: Tang, Y. Tom
; APPLICANT: Goodrich, Ryle
; APPLICANT: Liu, Chenghua
; APPLICANT: Ren, Feiyan
; APPLICANT: Wang, Dunrui
; APPLICANT: Drmanac, Radoje T.
; TITLE OF INVENTION: No. US20030219745A1e1 Nucleic Acids and
; TITLE OF INVENTION: Polypeptides
; FILE REFERENCE: 802CON
; CURRENT APPLICATION NUMBER: US/10/120,988
; CURRENT FILING DATE: 2002-04-11
; PRIOR APPLICATION NUMBER: 09/774,528
; PRIOR FILING DATE: 2001-01-30
; NUMBER OF SEQ ID NOS: 441
; SOFTWARE: pt_FL_genes Version 2.0
;
; SEQ ID NO 408
; LENGTH: 1068
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (100)..(1068)
US-10-120-988-408

Query Match   27.2%; Score 585.4; DB 16; Length 1068;
Best Local Similarity   95.9%; Pred. No. 1.7e-166;
Matches   601; Conservative   0; Mismatches 26; Indels   0; Gaps   0;

Qy   230 CCGGCCAGCCGCGAGCCGCGCCGCGCTCATGCCGAGGCCAGGAGGGATGG 289
             |||| ||||||| || | | || |||||||||||||
Db   44 CCGGTGCCCGCCGCTCCGGCGAAGCCGCGCCCGCCACTGCCCCAGAGCCAGAGCCAGAGGGATGG 103

Qy   290 TGGTAGTCACGGGGCGGGAGCCCAGACAGCCGTCGTCAGGACGTGCCATGTCCAGCTCTG 349
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   104 TGGTAGTCACGGGGCGGGAGCCCAGACAGCCGTCGTCAGGACGTGCCATGTCCAGCTCTG 163

Qy   350 ACGCCGAAGACGACTTTCTGGAGCCGGCCACGCGGCGGCGGCACGCGC 409
          ||||||||||||||||||||||||||||||||||||||||||||||
Db   164 ACGCCGAAGACGACTTTCTGGAGCCGGCCACGCGGCGGCGGCACGCGC 223

Qy   410 TGCCCCTGCTGCCACAGGAGTTTCCTGAGGTTGTTCCCCTTAACATCGGAGGGGCTCACT 469
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   224 TGCCCCTGCTGCCACAGGAGTTTCCTGAGGTTGTTCCCCTTAACATCGGAGGGGCTCACT 283

Qy   470 TCACTACACGCCTGTCCACACTGCGGTGCTACGAAGACACCATGTTGGCCATGTTCA 529
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   284 TCACTACACGCCTGTCCACACTGCGGTGCTACGAAGACACCATGTTGGCCATGTTCA 343

Qy   530 GTGGGCGGCACTACATCCCCACAGACTCCGAGGGCGGCGTG 589
          ||||||||||||||||||||||||||||||||||||||
Db   344 GTGGGCGGCACTACATCCCCACAGACTCCGAGGGCCGCCGTG 403

Qy   590 CACACTTTGGAGATGTGCTGAATTTCCTGCGCTCAGGGGACCTCGGGACTTCATGCCCT 649
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   404 CACACTTTGGAGATGTGCTGAATTTCCTGCGCTCAGGGGACCTCGGGACTTCATGCCCT 463

Qy   650 TTCGAGCTGTGTACAAAGAGGGCCCAGTACTTCATCGGGGACTTCATGCCCT 709
          |||||||||||||||||||||||||||||||||||||||||||||||||
Db   464 TTCGAGCTGTGTACAAAGAGGGCCCAGTACTTCATCGGGGACTTCATGCCCT 523

Qy   710 AGAACATGCAGCCACTGAAGGGCGAGAAGGTGCGCCAAGCGTTTCTGGGACTTCATGCCCT 769
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||

```
Db    524 AGAACATGCAGCCACTCGAAGGGCGAGAAGGTTCGCGCCAAGCGTTTCTGGGACTCATGCCCT 583
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||

Qy    770 ATTACAAAGACCACTTGGAGCCGGATTGTGGAGATCGCCCGGCTGCCGTCCAGCGGA 829
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||

Db    584 ATTACAAAGACCACTTGGAGCCGGATTGTGGAGATCGCCCGGCTGCCGTCCAGCGGA 643

Qy    830 AGGCCCCGCTTTGCCAAGCTCAAGAGCT 856
          ||||||||||||||||||||||||||||

Db    644 AGGCCCGCTTTGCCAAGCTCAAGGTCT 670
```

RESULT 4
US-10-296-115-18
; Sequence 18, Application US/10296115
; Publication No. US2004005324A1
; GENERAL INFORMATION:
; APPLICANT: Hyseq Inc
; TITLE OF INVENTION: No. US2004005324A1el Nucleic Acids and Polypeptides
; FILE REFERENCE: 784PCT
; CURRENT APPLICATION NUMBER: US/10/296,115
; CURRENT FILING DATE: 2002-11-18
; PRIOR APPLICATION NUMBER: US09/488,725
; PRIOR FILING DATE: 2000-01-21
; PRIOR APPLICATION NUMBER: US09/552,317
; PRIOR FILING DATE: 2000-04-25
; NUMBER OF SEQ ID NOS: 1478
; SEQ ID NO 18
; LENGTH: 519
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-296-115-18

```
Query Match              21.6%;  Score 465; DB 13; Length 519;
Best Local Similarity    94.2%;  Pred. No. 4.1e-130;
Matches    483;  Conservative    0;  Mismatches    0;  Indels   30;  Gaps    0;

Qy    224 CCCCCTCCGGCCAGCCCGCAGCGCCGCGGTCATGCCAGGGCTGCTCGGCGGAGCCAGCAG 283
          |||                        ||||||||||||||||||||||||||||||||||
Db      7 CCCGGGTCGACGATTTGCTCTCCGCCCGGAAGCCGCCCACTGCCCAGAGCCCAGAG 66

Qy    284 GGATGGTGGTAGTCACGGGGCGGGAGGCCAGAGACAGCAGCCGGTCGTCGCCAGCGGTGCCATGTCCA 343
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     67 GGATGGTGGTAGTCACGGGGCGGGAGGCCAGAGACAGCAGCCGGTCGTCGCCAGCGGTGCCATGTCCA 126

Qy    344 GCTCTGACGCCGAAGACGACTTTCTGGAGCCGGCCACGCGCGACGGCCAGGCGGGGC 403
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    127 GCTCTGACGCCGAAGACGACTTTCTGGAGCCGGCCACGCGCGACGGCCAGGCGGGGGC 186

Qy    404 ACGCGCTGCCCCCTGCTCCACAGGAGTTCCTGAGGTGTTGTTCCCCTTAACATCGGAGGGGG 463
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    187 ACGCGCTGCCCCCTGCTCCACAGGAGTTCCTGAGGTTGTTCCCCTTAACATCGGAGGGGG 246

Qy    464 CTCACTTCACTACACGGCCGCTGTCCACACTGCGGTGCTACGAAGACACCATGTTGGCAGCCA 523
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    247 CTCACTTCACTACACGGCCGCTGTCCACACTGCGGTGCTACGAAGACACCATGTTGGCAGCCA 306

Qy    524 TGTTCAGTGGGCGGCACTACATCCCCACAGACTCCGAGGGCCCGTACTTCATCGACCGAG 583
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    307 TGTTCAGTGGGCGGCACTACATCCCCACAGACTCCGAGGGCCCGTACTTCATCGACCGAG 366

Qy    584 ATGGCACACACTTTGGAGATGTGCTGAATTTCCTGCGCTCAGGGGACCTCCCACCCAGGG 643
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    367 ATGGCACACACTTTGGAGATGTGCTGAATTTCCTGCGCTCAGGGGACCTCCCACCCAGGG 426

Qy    644 AGGCTGTGTTCGAGCTGTGTACAAAGAGGGCCAGTACTATGGCCATCGGGCCCCTCCTGGAGC 703
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    427 AGCTGTGTTCGAGCTGTGTACAAAGAGGGCCAGTACTATGGCCATCGGGCCCCTCCTGGAGC 486

Qy    704 AGCTGGAGAACATGCCGCGCCACTGAAGGGCGAGA 736
          |||||||||||||||||||||||||||||||||||
Db    487 AGCTGGAGAACATGCCGCGCCACTGAAAGGCGAGA 519
```

RESULT 5
US-10-027-632-31358
; Sequence 31358, Application US/10027632
; Publication No. US2002019371A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; TITLE OF INVENTION: Polymorphisms in the Human Genome
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027,632
; CURRENT FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 31358
; LENGTH: 1100
; TYPE: DNA
; ORGANISM: Human
US-10-027-632-31358

```
Query Match              19.4%;  Score 417;  DB 13;  Length 1100;
Best Local Similarity    100.0%;  Pred. No. 2.5e-115;
Matches    417;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy    926 AAAGGGCCGGCTGTGTATTTGGAAACTCCCATTGGTCCAGAGACAGAACAATGAGAAGAAAT 985
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    684 AAAGGGCCGGCTGTGTATTTGGAAACTCCCATTGGTCCAGAGACAGAACAATGAGAAGAAAT 743

Qy    986 CCCCTGTCCAGTTGCCTGCAGGAGTATTCCAACACTTCATGGGCTAGAGGATTCCATTGA 1045
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    744 CCCCTGTCCAGTTGCCTGCAGGAGTATTCCAACACTTCATGGGCTAGAGGATTCCATTGA 803

Qy   1046 GATGGGGTTTACGTCTTGATTTTGAACACCTGTCAGCACTGTTCTCTGTTTGCATGGCAA 1105
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    804 GATGGGGTTTACGTCTTGATTTTGAACACCTGTCAGCACTGTTCTCTGTTTGCATGGCAA 863

Qy   1106 TTCTGACCCTTTTATGGCAACAACACCCCTGGACAACACCCAGATTGTAGATTGAGATCC 1165
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    864 TTCTGACCCTTTTATGGCAACAACACACCCCTGGACAACACCCAGATTGTAGATTGAGATCC 923

Qy   1166 AAAGGTAGAATTTCCAGACAGTCCAACCAAGGTATCAAGTGATGTTTCCAGAGTGGAAGG 1225
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    924 AAAGGTAGAATTTCCAGACAGTCCAACCAAGGTATCAAGTGATGTTTCCAGAGTGGAAGG 983

Qy   1226 CTCTCACCGTGTCCAGGATTTCTGGGGTTGTGTAAGCAGTACTACTGGCCATTGTGACCCTG 1285
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    984 CTCTCACCGTGTCCAGGATTTCTGGGGTTGTGTAAGCAGTACTACTGGCCATTGTGACCCTG 1043

Qy   1286 TTTTTACCTAATCATTCGTCTTTTAGGACATGGTTTTACCCGATCCTGGCAAA 1342
          |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1044 TTTTTACCTAATCATTCGTCTTTTAGGACATGGTTTTACCCGATCCTGGCAAA 1100
```

RESULT 6
US-10-027-632-31358
; Sequence 31358, Application US/10027632
; Publication No. US20030204075A9
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.

```
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; TITLE OF INVENTION: Polymorphisms in the Human Genome
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027,632
; CURRENT FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 31358
; LENGTH: 1100
; TYPE: DNA
; ORGANISM: Human
US-10-027-632-31358

Query Match             19.4%;  Score 417;  DB 16;  Length 1100;
Best Local Similarity   100.0%;  Pred. No. 2.5e-115;
Matches 417;  Conservative 0;  Mismatches 0;  Indels 0;  Gaps 0;

Qy    926 AAAGGCGGCTGTATTTGGAAACTCCCATTGGTCCAGAGAGACAGAACAATGAGAGAGAAAT 985
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    684 AAAGGCGGCTGTATTTGGAAACTCCCATTGGTCCAGAGAGACAGAACAATGAGAGAGAAAT 743

Qy    986 CCCCTGTCCAGTTGCCTGCAGGAGTATTCCAACACTTCAATGGGCTAGAGGATTCCATTGA 1045
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    744 CCCCTGTCCAGTTGCCTGCAGGAGTATTCCAACACTTCAATGGGCTAGAGGATTCCATTGA 803

Qy   1046 GATGGGGTTTACGTCTTGATTTTGAACACCTGTCAGCACCTGTTCTCGTTTGCATGGCAA 1105
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    804 GATGGGGTTTACGTCTTGATTTTGAACACCTGTCAGCACCTGTTCTCGTTTGCATGGCAA 863

Qy   1106 TTCTGACCCTTTATGGCAACACACCCCTGGGACAACAACCCAGATTTGTAGATTGAGATCC 1165
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    864 TTCTGACCCTTTATGGCAACACACCCCTGGGACAACAACCCAGATTTGTAGATTGAGATCC 923

Qy   1166 AAAGGTAGAATTTCCAGACAGTCCAACCAAGGTATCAAGTGATGTTTCCAGAGTGGAAGG 1225
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    924 AAAGGTAGAATTTCCAGACAGTCCAACCAAGGTATCAAGTGATGTTTCCAGAGTGGAAGG 983

Qy   1226 CTCTCACCGTGTCCCAGGATTTCTGCGGGTTTGTAAGCAGTACTGGCCATTTGTGACCCTG 1285
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    984 CTCTCACCGTGTCCCAGGATTTCTGCGGGTTTGTAAGCAGTACTGGCCATTTGTGACCCTG 1043

Qy   1286 TTTTTTACCTAATCATTCTGTCTTTTTAGGACATGGTTTTACCCGATCCCTGGCAAA 1342
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1044 TTTTTTACCTAATCATTCTGTCTTTTTAGGACATGGTTTTACCCGATCCCTGGCAAA 1100

RESULT 7
US-10-024-579-8
; Sequence 8, Application US/10024579
; Publication No. US20020119522A1
; GENERAL INFORMATION:
; APPLICANT: Friddle, Carl Johan
; APPLICANT: Gerhardt, Brenda
; APPLICANT: Hilbun, Erin
; APPLICANT: Turner, C. Alexander Jr.
; TITLE OF INVENTION: No. US20020119522A1e1 Human Ion Channel-Related Proteins
; TITLE OF INVENTION: and Polynucleotides Encoding the Same
; FILE REFERENCE: LEX-0274-USA
; CURRENT APPLICATION NUMBER: US/10/024,579
```

```
; CURRENT FILING DATE: 2001-12-18
; PRIOR APPLICATION NUMBER: US 60/258,595
; PRIOR FILING DATE: 2000-12-28
; NUMBER OF SEQ ID NOS: 17
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 8
; LENGTH: 680
; TYPE: DNA
; ORGANISM: homo sapiens
US-10-024-579-8

Query Match             18.0%;  Score 387.8;  DB 14;  Length 680;
Best Local Similarity   88.2%;  Pred. No. 1.4e-106;
Matches 456;  Conservative 1;  Mismatches 3;  Indels 57;  Gaps 1;

Qy    146 CGGGTCAGGCCCCAGCTGGGCGGGTGGGCGAGCGGGTCGGCGGTTGAGGGAGGCCACGCCCTCCCGC 205
          ||||||||||||||||||||||||||||||||||||||||||||||||
Db      1 CGGGTCAGGCCCCAGCTGGGCGGGTGGGCGAGCGGGTCGGCGGTTGAGGGAGGCCACGCCCTCCCGC 60

Qy    206 CTGCGCACTGCCTCTCGCCCCCTCCGGCCAGCCCGCAGCCCGGCCGGCTCATGCCAGGCG 265
          |||||||||||||||||||||||||||||||||||||||||||| |||||||||||||
Db     61 CTGCGCACTGCCTCTCGCCCCCTCCGGCCAGCCCGCAGCCCGGCCGGCGTSATGCCAGGCG 120

Qy    266 CTGCTCGGC------------------------------------------------------ 274
          |||||||||
Db    121 CTGCTCGGCCGGTAGGGAGTGCCGGCGGGAGCGGGAGCGGGAGCGGGAGCGGGAGCGGG 180

Qy    275 ------GAGCCAGAGCCAGAGAGGGATGGTGGTAGTCAGGGAGGCGGGAGCAGCGCCGTCGTCAGG 328
                ||||||||||||||||||||||||||||||||||||||||||||||||||
Db    181 TGCCCAGAGCCAGAGCCAGAGAGGGATGGTGGTAGTCAGGGAGGCGGGAGCAGCGCCGTCGTCAGG 240

Qy    329 ACGGTGCCATGTCCAGCTCTGACGCGCGAAGCGACGTTTCTGGAGCGCGGCCACGCCGACGG 388
          |||||| |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    241 ACGGTCCATGTCCAGCTCTGACGCGCGAAGCGACGTTTCTGGAGCGCGGCCACGCCGACGG 300

Qy    389 CCACCGCAGGGCGGGGCACGCGCTGCCCCTGCGTGCCACAGGAGTGTTCCTGAGGTTGTTCCCC 448
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    301 CCACCGCAGGGCGGGGCACGCGCTGCCCCTGCGTGCCACAGGAGTGTTCCTGAGGTTGTTCCCC 360

Qy    449 TTAACATCGGAGGGCTCACTCACTACACCCGTGTCCACACGCCGTGCTACGAAGACA 508
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    361 TTAACATCGGAGGGCTCACTCACTACACCCGTGTCCACACGCCGTGCTACGAAGACA 420

Qy    509 CCATGTTGGCAGCCATGTTCAGTGGCGGCACTACATCCCACACTCCGAGGGCCGGT 568
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    421 CCATGTTGGCAGCCATGTTCAGTGGCGGCACTACATCCCACACTCCGAGGGCCGGT 480

Qy    569 ACTTCATCGACCGAGATGGCACACACTTTGGAGATGT 605
          |||||||||||||||||||||||||||||||||
Db    481 ACTTCATCGACCGAGATGGCACACACTTTGGGTATGT 517

RESULT 8
US-10-024-579-4
; Sequence 4, Application US/10024579
; Publication No. US20020119522A1
; GENERAL INFORMATION:
; APPLICANT: Friddle, Carl Johan
; APPLICANT: Gerhardt, Brenda
; APPLICANT: Hilbun, Erin
; APPLICANT: Turner, C. Alexander Jr.
; TITLE OF INVENTION: No. US20020119522A1e1 Human Ion Channel-Related Proteins
; TITLE OF INVENTION: and Polynucleotides Encoding the Same
; FILE REFERENCE: LEX-0274-USA
; CURRENT APPLICATION NUMBER: US/10/024,579
; CURRENT FILING DATE: 2001-12-18
; PRIOR APPLICATION NUMBER: US 60/258,595
; PRIOR FILING DATE: 2000-12-28
; NUMBER OF SEQ ID NOS: 17
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 4
; LENGTH: 363
; TYPE: DNA
```

```
; ORGANISM: homo sapiens
US-10-024-579-4

Query Match      14.6%;     Score 315.2; DB 14; Length 363;
Best Local Similarity   99.1%;   Pred. No. 1e-84;
Matches  317;   Conservative  0;   Mismatches  0;   Indels  3;   Gaps  0;

Qy  286  ATGGTGGTAGTCACGGGCGGAGCCAGACAGCCGTCGTCAGGACGGTGCCATGTCCAGC 345
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1  ATGGTGGTAGTCACGGGCGGAGCCAGACAGCCGTCGTCAGGACGGTGCCATGTCCAGC  60

Qy  346  TCTGACGCCGAAGACGACTTTCTGGAGCCGCCACGCGGCCGCCACGCAGCGGCGGGCAC 405
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   61  TCTGACGCCGAAGACGACTTTCTGGAGCCGCCACGCGGCCGCCACGCAGCGGCGGGCAC 120

Qy  406  GCGCTGCCCCTGCTGCCACAGGAGTTTCCTGAGGTTGTTCCCTTAACATCGGAGGGGCT 465
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  121  GGCGCTGCCCCTGCTGCCACAGGAGTTTCCTGAGGTTGTTCCCTTAACATCGGAGGGGCT 180

Qy  466  CACTTCACTACACGCCTGTCCACACTGCGGTGCTACGAAGACACCATGTTGGCAGCCATG 525
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  181  CACTTCACTACACGCCTGTCCACACTGCGGTGCTACGAAGACACCATGTTGGCAGCCATG 240

Qy  526  TTCAGTGGGCGGCACTACATCCCCACAGACTCCGAGGGCCGGTACTTCATCGACCGAGAT 585
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  241  TTCAGTGGGCGGCACTACATCCCCACAGACTCCGAGGGCCGGTACTTCATCGACCGAGAT 300

Qy  586  GGACACACTTTGGAGATGT 605
         ||||||||||||| ||||
Db  301  GGCACACTTTGGGTATGT 320
```

RESULT 9
US-09-864-761-1385/c
; Sequence 1385, Application US/09864761
; Patent No. US20020048763A1
GENERAL INFORMATION:
  APPLICANT: Penn, Sharron G.
  APPLICANT: Rank, David R.
  APPLICANT: Hanzel, David K.
  APPLICANT: Chen, Wensheng
  TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
  TITLE OF INVENTION: GENE EXPRESSION ANALYSIS BY MICROARRAY
  FILE REFERENCE: Aeomica-X-1
  CURRENT APPLICATION NUMBER: US/09/864,761
  CURRENT FILING DATE: 2001-05-23
  PRIOR APPLICATION NUMBER: US 60/180,312
  PRIOR FILING DATE: 2000-02-04
  PRIOR APPLICATION NUMBER: US 60/207,456
  PRIOR FILING DATE: 2000-05-26
  PRIOR APPLICATION NUMBER: US 09/632,366
  PRIOR FILING DATE: 2000-08-03
  PRIOR APPLICATION NUMBER: GB 24263.6
  PRIOR FILING DATE: 2000-10-04
  PRIOR APPLICATION NUMBER: US 60/236,359
  PRIOR FILING DATE: 2000-09-27
  PRIOR APPLICATION NUMBER: PCT/US01/00666
  PRIOR FILING DATE: 2001-01-30
  PRIOR APPLICATION NUMBER: PCT/US01/00667
  PRIOR FILING DATE: 2001-01-30
  PRIOR APPLICATION NUMBER: PCT/US01/00664
  PRIOR FILING DATE: 2001-01-30
  PRIOR APPLICATION NUMBER: PCT/US01/00669
  PRIOR FILING DATE: 2001-01-30
  PRIOR APPLICATION NUMBER: PCT/US01/00665
  PRIOR FILING DATE: 2001-01-30
  PRIOR APPLICATION NUMBER: PCT/US01/00668
  PRIOR FILING DATE: 2001-01-30
  PRIOR APPLICATION NUMBER: PCT/US01/00663
  PRIOR FILING DATE: 2001-01-30
  PRIOR APPLICATION NUMBER: PCT/US01/00662
  PRIOR FILING DATE: 2001-01-30
  PRIOR APPLICATION NUMBER: PCT/US01/00661

  PRIOR FILING DATE: 2001-01-30
  PRIOR APPLICATION NUMBER: PCT/US01/00670
  PRIOR FILING DATE: 2001-01-30
  PRIOR APPLICATION NUMBER: US 60/234,687
  PRIOR FILING DATE: 2000-09-21
  PRIOR APPLICATION NUMBER: US 09/608,408
  PRIOR FILING DATE: 2000-06-30
  PRIOR APPLICATION NUMBER: US 09/774,203
  PRIOR FILING DATE: 2001-01-29
  NUMBER OF SEQ ID NOS: 49117
  SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
SEQ ID NO 1385
  LENGTH: 473
  TYPE: DNA
  ORGANISM: Homo sapiens
  FEATURE:
  OTHER INFORMATION: MAP TO AC006001.2
  OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 4.6
  OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 12
  OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 5.4
  OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 4.8
  OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 7.3
  OTHER INFORMATION: EXPRESSED IN BT474, SIGNAL = 14
  OTHER INFORMATION: EXPRESSED IN HBL100, SIGNAL = 8.5
  OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 5.5
  OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 2.4
  OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 6.4
US-09-864-761-1385

```
Query Match      8.1%;     Score 174; DB 9; Length 473;
Best Local Similarity   83.2%;   Pred. No. 9.7e-42;
Matches  198;   Conservative  0;   Mismatches  40;   Indels  0;   Gaps  0;

Qy  368  TGGAGCCGGCCACGCCAGCGCCAGCGGGGCGCACGCGCTGCCCCTGCTGCCACAGG 427
         |||||  ||||| ||||   |||||||||||||  |||||||||||||||||||
Db  461  TGGAGCAGCCCCAGCTCTCATTCCCGTTGCCTGCTGGAGAGGGTCACTTCACTACACGCCTGTGGTGATTTCTTCC 402

Qy  428  AGTTTCCTGAGGTTGTTCCCCTTAACATCGGAGGGGCTCACTTCACTACACGCCTGTCCA 487
         |||||||||||||||||||  |||||||||||||||||||||||||||||||||||||||
Db  401  AGTTTCCTGAGGTTGTTCCCCTTAACATCGGAGGGGCTCACTTCACTACACGCCTGTCCA 342

Qy  488  CACTGCGGTGCTACGAAGACACCATGTTGGCAGCCATGTTCAGTGGGCGGCACTACATCC 547
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  341  CACTGCGGTGCTACGAAGACACCATGTTGGCAGCCATGTTCAGTGGGCGGCACTACATCC 282

Qy  548  CCACAGACTCCGAGGGCCGGTACTTCATCGACCGAGATGGCACACACTTTGGAGATGT 605
         ||||| ||||||||||||||||||||||||||||||||||||||||||||||| |||
Db  281  CCACGGACTCCGAGGGCCGGTACTTCATCGACCGAGATGGCACACACTTTGGGTATGT 224
```

RESULT 10
US-10-086-156-29
; Sequence 29, Application US/10086156
; Publication No. US20030054989A1
GENERAL INFORMATION:
  APPLICANT: Bristol-Myers Squibb Company
  TITLE OF INVENTION: POLYNUCLEOTIDE ENCODING TWO NOVEL HUMAN POTASSIUM CHANNEL BETA-SU
  TITLE OF INVENTION: K+betaM4 and K+betaM5
  FILE REFERENCE: D0115NP
  CURRENT APPLICATION NUMBER: US/10/086,156
  CURRENT FILING DATE: 2002-02-28
  PRIOR APPLICATION NUMBER: US 60/272,190
  PRIOR FILING DATE: 2001-02-28
  PRIOR APPLICATION NUMBER: US 60/274,258
  PRIOR FILING DATE: 2001-03-07
  NUMBER OF SEQ ID NOS: 98
  SOFTWARE: PatentIn version 3.0
SEQ ID NO 29
  LENGTH: 583
  TYPE: DNA
  ORGANISM: homo sapiens
US-10-086-156-29

Query Match   8.1%; Score 174; DB 15; Length 583;
Best Local Similarity 83.2%; Pred. No. 1.1e-41;
Matches 198; Conservative 0; Mismatches 40; Indels 0; Gaps 0;

```
Qy  368 TGGAGCCGGCCACGCCGACGGCCACGCAGGCGGGCACGCGGCTGCCCCTGCTGCCACAGG 427
            |||| ||| |  ||
Db  145 TGGAGCAGCCCCAGCTCTCATTCCCCGTTGCCTGCTGAGAGCCCTGGTGATTTCTTTCC 204

Qy  428 AGTTTCCTGAGGTTGTTCCCCTTAACATCGGAGGGGCTCACTTCACTACACGCCTGTCCA 487
            |||||||||||||||||||||||||||||||||||||||||||||||||
Db  205 AGTTTCCTGAGGTTGTTCCCCTTAACATCGGAGGGGCTCACTTCACTACACGCCCTGTCCA 264

Qy  488 CACTGCGGTGCTACGAAGACACCATGTTGGCAGCACCATGTTCAGTGGGCGGCCACTACATCC 547
            ||||  |||||||| ||||||||
Db  265 CACTGCGGTGCTACGAAGACACCATGTTGGCAGCACCATGTTGGCAGCCATGTTCAGTGGGCGGCGGCCACTACATCC 324

Qy  548 CCACAGACTCCGAGGGCCGGTACTTCATCGACCGAGATGGCACACACTTTGGAGATGT 605
            |||||||||||||||||||||||||||||||||||||||
Db  325 CCACGGACTCCGAGGGCCGGTACTTCATCGACCGAGATGGCACACACTTTGGGTATGT 382
```

RESULT 11
US-10-024-579-6
; Sequence 6, Application US/10024579
; Publication No. US20020119522A1
; GENERAL INFORMATION:
; APPLICANT: Friddle, Carl Johan
; APPLICANT: Gerhardt, Brenda
; APPLICANT: Hilbun, Erin
; APPLICANT: Turner, C. Alexander Jr.
; TITLE OF INVENTION: No. US2002011 9522A1e1 Human Ion Channel-Related Proteins
; TITLE OF INVENTION: and Polynucleotides Encoding the Same
; FILE REFERENCE: LEX-0274-USA
; CURRENT APPLICATION NUMBER: US/10/024,579
; CURRENT FILING DATE: 2001-12-18
; PRIOR APPLICATION NUMBER: US 60/258,595
; PRIOR FILING DATE: 2000-12-28
; NUMBER OF SEQ ID NOS: 17
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 6
; LENGTH: 321
; TYPE: DNA
; ORGANISM: homo sapiens
US-10-024-579-6

Query Match   8.0%; Score 172.4; DB 14; Length 321;
Best Local Similarity 82.8%; Pred. No. 2.3e-41;
Matches 197; Conservative 0; Mismatches 41; Indels 0; Gaps 0;

```
Qy  368 TGGAGCCGGCCACGCCGACGGCCACGCAGGCGGGCACGCGGCTGCCCCTGCTGCCACAGG 427
            |||| ||| |  ||
Db   41 TGGAGCAGCCCCAGCTCTCATTCCCCGTTGCCTGCTGAGAGCCCTGGTGATTTCTTTCC 100

Qy  428 AGTTTCCTGAGGTTGTTCCCCTTAACATCGGAGGGGCTCACTTCACTACACGCCTGTCCA 487
            |||||||||||||||||||||||||||||||||||||||||||||||||
Db  101 AGTTTCCTGAGGTTGTTCCCCTTAACATCGGAGGGGCTCACTTCACTACACGCCCTGTCCA 160

Qy  488 CACTGCGGTGCTACGAAGACACCATGTTGGCAGCCATGTTCAGTGGGCGGCCACTACATCC 547
            ||||  |||||||||||||||||
Db  161 CACTGCGGTGCTACGAAGACACCATGTTGGCATCCATGTTGGCATCCATGTTCAGTGGGCGGCGGCCACTACATCC 220

Qy  548 CCACAGACTCCGAGGGCCGGTACTTCATCGACCGAGATGGCACACACTTTGGAGATGT 605
            |||||||||||||||||||||||||||||||||||||||
Db  221 CCACGGACTCCGAGGGCCGGTACTTCATCGACCGAGATGGCACACACTTTGGGTATGT 278
```

RESULT 12
US-09-864-761-18144/c
; Sequence 18144, Application US/09864761
; Patent No. US20020048763A1
; GENERAL INFORMATION:
; APPLICANT: Penn, Sharron G.
; APPLICANT: Rank, David R.

; APPLICANT: Hanzel, David K.
; APPLICANT: Chen, Wensheng
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
; TITLE OF INVENTION: GENE EXPRESSION ANALYSIS BY MICROARRAY
; FILE REFERENCE: Aeomica-X-1
; CURRENT APPLICATION NUMBER: US/09/864,761
; CURRENT FILING DATE: 2001-05-23
; PRIOR APPLICATION NUMBER: US 60/180,312
; PRIOR FILING DATE: 2000-02-04
; PRIOR APPLICATION NUMBER: US 60/207,456
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: US 09/632,366
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: GB 24263.6
; PRIOR FILING DATE: 2000-10-04
; PRIOR APPLICATION NUMBER: US 60/236,359
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: PCT/US01/00666
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00667
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00664
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00669
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00665
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00668
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00663
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00662
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00661
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: US 60/234,687
; PRIOR FILING DATE: 2000-09-21
; PRIOR APPLICATION NUMBER: US 09/608,408
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: US 09/774,203
; PRIOR FILING DATE: 2001-01-29
; PRIOR APPLICATION NUMBER: PCT/US01/00670
; PRIOR FILING DATE: 2001-01-30
; NUMBER OF SEQ ID NOS: 49117
; SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
SEQ ID NO 18144
; LENGTH: 173
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: MAP TO AC006001.2
; OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 4.6
; OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 12
; OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 5.4
; OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 4.8
; OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 7.3
; OTHER INFORMATION: EXPRESSED IN BT474, SIGNAL = 14
; OTHER INFORMATION: EXPRESSED IN HBL100, SIGNAL = 8.5
; OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 5.5
; OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 2.4
; OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 6.4
; OTHER INFORMATION: EST HUMAN HIT: AI674184.1, EVALUE 5.00e-93
; OTHER INFORMATION: NT HIT: AJ001455.1, EVALUE 1.60e+00
; OTHER INFORMATION: SWISSPROT HIT: Q13829, EVALUE 1.00e-06
US-09-864-761-18144

Query Match   7.9%; Score 170.4; DB 9; Length 173;
Best Local Similarity 99.4%; Pred. No. 6.3e-41;
Matches 171; Conservative 0; Mismatches 1; Indels 1; Gaps 0;

```
Qy  428 AGTTTCCTGAGGTTGTTCCCCTTAACATCGGAGGGGCTCACTTCACTACACGCCTGTCCA 487
            ||||||||||||||||||||||||||||||||||||||||||||||||||
Db  172 AGTTTCCTGAGGTTGTTCCCCTTAACATCGGAGGGGCTCACTTCACTACACGCCTGTCCA 113
```

```
Qy      488  CACTGCGGTGCTACGAAGACACCATGTTGGCAGCCATGTTCAGTGGGGCGGCACTACATCC  547
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      112  CACTGCGGTGCTACGAAGACACCATGTTGGCAGCCATGTTCAGTGGGGCGGCACTACATCC  53

Qy      548  CCACGAGACTCCGAGGGCCCGTACTTCATCGACCGAGATGGCACACACTTTGG  599
             |||| |||||||||||||||||||||||||||||||||||||||||||||||
Db       52  CCACGGACTCCGAGGGCCCGTACTTCATCGACCGAGATGGCACACACTTTGG  1
```

RESULT 13
US-10-240-425-1575
; Sequence 1575, Application US/10240425
; Publication No. US2004003502A1
; GENERAL INFORMATION:
;   APPLICANT: Williams, Amanda
;   APPLICANT: Boland, Joseph F.
;   APPLICANT: Lord, Reginald V.
;   APPLICANT: Alvarez, Chris
;   APPLICANT: Wetzel, Jon C.
;   APPLICANT: Scherf, Uwe
;   APPLICANT: Vockley, Joseph G.
;   TITLE OF INVENTION: Gene Expression Profiles in Esophageal Tissue
;   FILE REFERENCE: 44921-5026
;   CURRENT APPLICATION NUMBER: US/10/240,425
;   CURRENT FILING DATE: 2002-09-30
;   PRIOR APPLICATION NUMBER: PCT/US01/09847
;   PRIOR FILING DATE: 2001-03-28
;   PRIOR APPLICATION NUMBER: US 60/193,446
;   PRIOR FILING DATE: 2000-03-31
;   NUMBER OF SEQ ID NOS: 1588
;   SOFTWARE: PatentIn Ver. 2.1
US-10-240-425-1575
;   SEQ ID NO 1575
;   LENGTH: 18534
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   OTHER INFORMATION: Genbank Accesion No. US20040033502A1
US-10-240-425-1575

Query Match   5.4%; Score 116.4; DB 13; Length 18534;
Best Local Similarity   75.8%; Pred. No. 3.2e-23;
Matches   144; Conservative   0; Mismatches   46; Indels   0; Gaps   0;

```
Qy     1965  TATCTGACATGTAAGACCAGCCTACACATTGGGGTGGTGCGAGGGGCTCACACTTGTAAT  2024
             |||| ||| ||| || |||||||||||| |||||||| | | ||||| | ||||||||||
Db     2188  TATCAGACTTGAGATATAAAATGTCAGGCTGAGCGGATGCAGTGGCTTATACCTGTAAT  2247

Qy     2025  CCTAGCACTTTGGAAGGCTGAGGTGGGCAGATTGCTTGAGCACAGGAGTTCCAGACCAGC  2084
             || |||||| | ||||||| |||||||||||| ||| |||||| | ||||||||||||||
Db     2248  CCCAACATTTTGGGAGGCCGAAGTGGGTCGATCCGATCGCTTGAGTCCAGGAGTTTCAGACCAGC  2307

Qy     2085  CTGAGCAACATGCGAAATCCTGTCTCTTCAAGAAAATAAAATAATAATAATAAAAA  2144
             |||||||||||||||||||||||||||| | || || |||||||||||||||||||
Db     2308  CTGAGCAACATTGGGAAACCCTGTCTCTCTCACAAAAAAAAAAAAAAAAAAAAAAAA  2367

Qy     2145  AAAAAAAAA  2154
             |||||||||
Db     2368  AAAAAAAAA  2377
```

RESULT 14
US-10-087-192-1732
; Sequence 1732, Application US/10087192
; Publication No. US2002018258A1
; GENERAL INFORMATION:
;   APPLICANT: Morris, David W.
;   APPLICANT: Engelhard, Eric K.
;   TITLE OF INVENTION: NOVEL COMPOSITIONS AND METHODS FOR
;   TITLE OF INVENTION: CANCER
;   FILE REFERENCE: 5294S2000122
;   CURRENT APPLICATION NUMBER: US/10/087,192
;   CURRENT FILING DATE: 2002-03-01

; PRIOR APPLICATION NUMBER: US 09/747,377
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: US 09/798,586
; PRIOR FILING DATE: 2001-03-02
; NUMBER OF SEQ ID NOS: 2059
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 1732
;   LENGTH: 60430
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-10-087-192-1732

Query Match   5.4%; Score 116.4; DB 13; Length 60430;
Best Local Similarity   75.8%; Pred. No. 6.7e-23;
Matches   144; Conservative   0; Mismatches   46; Indels   0; Gaps   0;

```
Qy     1965  TATCTGACATGTAAGACCAGCCTACACATTGGGGTGGTGCGAGGGGCTCACACTTGTAAT  2024
             |||| ||| ||| || |||||||||||| |||||||| | | ||||| | ||||||||||
Db     7741  TATCAGACTTGAGATATAAAATGTCAGGCTGAGCGGATGCAGTGGCTTATACCTGTAAT  7800

Qy     2025  CCTAGCACTTTGGAAGGCTGAGGTGGGCAGATTGCTTGAGCACAGGAGTTCCAGACCAGC  2084
             || |||||| | ||||||| |||||||||||| ||| |||||| | ||||||||||||||
Db     7801  CCCAACATTTTGGGAGGCCGAAGTGGGTCGATCGGTTGAGTCCAGGAGTTTCAGACCAGC  7860

Qy     2085  CTGAGCAACATGCGAAATCCTGTCTCTTCAAGAAAATAAAATAATAATAATAAAAA  2144
             |||||||||||||||||||||||||||| | || || |||||||||||||||||||
Db     7861  CTGAGCAACATTGGGAAACCCTGTCTCTCTCACAAAAAAAAAAAAAAAAAAAAAAAA  7920

Qy     2145  AAAAAAAAA  2154
             |||||||||
Db     7921  AAAAAAAAA  7930
```

RESULT 15
US-10-027-632-110966/c
; Sequence 110966, Application US/10027632
; Publication No. US2002019837A1
; GENERAL INFORMATION:
;   APPLICANT: Wang, David G.
;   TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
;   TITLE OF INVENTION: Polymorphisms in the Human Genome
;   FILE REFERENCE: 108827.129
;   CURRENT APPLICATION NUMBER: US/10/027,632
;   CURRENT FILING DATE: 2002-04-30
;   PRIOR APPLICATION NUMBER: US 60/218,006
;   PRIOR FILING DATE: 2000-07-12
;   PRIOR APPLICATION NUMBER: US 60/198,676
;   PRIOR FILING DATE: 2000-04-20
;   PRIOR APPLICATION NUMBER: US 60/193,483
;   PRIOR FILING DATE: 2000-03-29
;   PRIOR APPLICATION NUMBER: US 60/185,218
;   PRIOR FILING DATE: 2000-02-24
;   PRIOR APPLICATION NUMBER: US 60/167,363
;   PRIOR FILING DATE: 1999-11-23
;   PRIOR APPLICATION NUMBER: US 60/156,358
;   PRIOR FILING DATE: 1999-09-28
;   PRIOR APPLICATION NUMBER: US 60/146,002
;   PRIOR FILING DATE: 1999-08-09
;   NUMBER OF SEQ ID NOS: 325720
;   SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 110966
;   LENGTH: 646
;   TYPE: DNA
;   ORGANISM: Human
US-10-027-632-110966

Query Match   5.4%; Score 116.2; DB 13; Length 646;
Best Local Similarity   71.0%; Pred. No. 4.3e-24;
Matches   154; Conservative   0; Mismatches   63; Indels   0; Gaps   0;

```
Qy     1931  CCATTCAAAATAACATGCATTCTTGGAACCTTGTATCTGACATGTAAGACCAGCCTACA  1990
             |||  |||| || || | |||| ||| |||| ||| || ||  | ||| ||||  || |
Db      276  CCTGTAAATACTAAGTTTGATCATCAGTGACTAAGGTGATCTTCCAGGTCTCCCATTGTAAT  217
```

```
Qy   1991 CATTGGGGTGGGTGCAGGGGCTCACACTTGTAATCCTAGCACTTTGGAAGGCTGAGGTGG 2050
          ||||  |  |  ||  |||||  ||||| | |||||||||||| |||||||||| ||||||
Db    216 TATTTTGCTGGCTACAGTGCGCTCCCACTTGTAATCCCAGCACTTTGGAAGGCCGAGGTGG  157

Qy   2051 GCAGATTGCTTGAGCACAGAGTTCCAGACCAGCCTGAGCAACATGGCGAAATCCTGTCT 2110
          | ||||||||||||  |||||||||| |  ||||||||||||||| |||| |||||||
Db    156 GAGGATTGCTTGAGCCCAGGAGTTCTACACCAGCCTGAGCAACATGGGGAAACTCTGTCA  97

Qy   2111 CTTCAAGAAATAAAATAATAATAATAATAAAAAAAA 2147
          ||| |  ||||||||   || || ||||||||||||
Db     96 CTACTGAAAATACAAAAGTACAAAAAAAAAAAAAAA  60
```

Search completed: October   8, 2004, 12:45:27
Job time : 1032 secs